NEW SERIES, NO. 140 (VOL. XVIII) DECEMBER, 1922

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Formerly the Quarterly Publication of the American Statistical Association

THE MEASUREMENT OF DIFFERENCES BETWEEN VARIABLE QUANTITIES

By Franz Boas, *Columbia University*

In biological statistics it is often necessary to study the differences between two variable types. The problem may be exemplified by a consideration of the differences between the types represented by populations of various countries, as, for instance, between the populations of Sweden, Switzerland, and Central Africa. It is obvious that the type of Sweden differs less from that of Switzerland than the latter differs from the type of Central Africa. Nevertheless, it is difficult to say just what is meant by greater or lesser difference in type. The attempt to establish and describe varieties of races according to characteristic features that are considered as significant from a morphological point of view, suffers, therefore, from a lack of clarity of concept.

The differences between the averages of types have been utilized for the purpose of segregating subtypes of human races, as, for instance, in the classification of the local types into which the European race may be divided. Pigmentation, stature, form of hair, head, face, and nose have been so utilized. For example, local types have been described by Deniker [1] by assigning to each group peoples among whom certain average values of measurements are found. All those that have average statures, head indices, facial forms, nose forms, and pigmentation falling within certain limits that may be expressed numerically, were assigned by him to a certain subrace. Although it is possible to give in this manner a definite description of local types, the biological significance of the observed differences remains undetermined. Obviously the classification obtained by the method here indicated will vary according to the limits which we set for each division. If we call tall

[1] *The Races of Man.*

those populations whose average stature is more than 170 cm., their assignment to a subdivision will not be the same as the one obtained when we call those tall whose average stature is more than 172 cm. If no valid reason can be given for the choice of one or the other limit, then the subtype so established can have a conventional descriptive meaning only. If we wish to establish a biologically significant classification we should have to prove that the descriptive features selected are morphologically significant. Furthermore, it would be necessary to distinguish between environmental and hereditary influences that determine the particular features which are made the basis of the classification. As a matter of fact, this study has never been made; and since the lines of demarcation between classes are arbitrary, these classes will be only a convenient schematic review of the distribution of certain selected combinations of descriptive features.[1]

In the following pages I wish to discuss the question whether a valid method of comparing closely allied forms can be found, so that arbitrary classifications may be avoided and measurable differences between types established.

It may seem that maps showing the distribution of a single feature or of combined features would give this information. Retzius' maps of Sweden, Livi's maps of Italy, Virchow's map of hair color in Germany, anthropological maps of France, England, and Spain, all illustrate the distribution of forms of the body, either by showing the areas in which the same average value of a measurement occurs, or areas in which certain selected values occur with equal frequency. The maps are intended to convey the impression that sameness of average values or of frequencies indicate the occurrence of the same racial forms. They also indicate that the differences between types are equivalent whenever the differences between averages or between frequencies of occurrence of selected values are the same. This has often led to the interpretation that the values whose frequency is shown represent separate racial types. Thus, the frequency of long-headedness in an area is often said to mean that a long-headed race forms a certain proportion of the population, although no biological basis can be given for the claim that the arbitrarily selected values represent a separate racial type.

The essential difficulty of our problem may be made clearer by the following considerations. Each racial type is variable. When we study the distribution of any particular feature, let us say of the cephalic index among European types, we find that the forms which occur in each area are variable, and the individuals composing the populations of

[1] See also St. Poniatowski, *Ueber den Wert der Index Klassifikation*, Archiv. für Anthropologie, N. F. Vol. 10, p. 1.

different areas show in part the same numerical values of the measurement. The distribution of forms in each population is such that the types overlap. The average cephalic index in Sweden may be 77; in Bavaria 85. Nevertheless, there will be many individuals that have the index 81, both in Bavaria and in Sweden; and according to this particular feature individuals may belong to either group. We know that if two regions are not too far apart, it is quite impossible to assign with certainty a single individual to either of them.

If we assume for the moment that the variability found both in Sweden and in Bavaria were very low, so that the highest cephalic index occurring among Swedes would be not more than 80 and the lowest occurring in Bavaria not less than 82, then the two series would appear to us entirely distinct. It would be quite inadmissible to claim that the differences between the two pairs of groups were the same in the cases of greater variability (which has actually been observed) and the lesser variability (which has here been assumed), although in both cases the averages show the same differences. In the latter case we judge that the difference is greater, or perhaps better, more fundamental.

Obviously our judgment is influenced by the degree of variability, still more, by the degree of overlapping of the two series. Only if we assume quite arbitrarily that the individuals that show the average values of the measurement in question—or some other selected value—were the true representatives of the whole population, and that all others were present only as foreign, intrusive elements, or if their occurrence represented modifications of the typical form due to extraneous causes—only under these conditions could we say that the difference between the selected values represents the difference between the types. A concept of variability like the one involved in these assumptions is, however, quite inadmissible. The group must be considered as a *class* and its variability be determined by the definition of the class in question. Our detailed study of the class will always be directed toward the discovery of new principles of classification by means of which subclasses are formed whose variability will be less than that of the original class. In this way we try to define the newly formed subclasses more sharply than the original class, and the advance in our knowledge consists in the discovery of the factors that make the subclasses more determinate. It would be quite arbitrary to select one particular individual as the type, and to claim either that all others are not really members of the class or that they are modified forms of the type. This method of procedure would contravene the fundamental concept of variability, for a variable comprises *all* the representatives of a class, the individual components of which are only defined in so far

as they are members of the class—this in contradistinction to constants which are assumed to be completely defined and must therefore be the same in every case.

As soon as these principles are held clearly in mind, it appears that the ordinary definition of arithmetical difference is not applicable in our case. The term "difference" as applied to variables does not mean the same as the term "difference" applied to constants. Variables cannot be brought into a measurable series by the same means that we use for constants which may be compared by means of an arbitrary standard that is also constant.

The problem before us is how to overcome these difficulties—how to give a definite meaning to the differences between variables and make these differences measurable.

The question has been treated by G. H. Mollison [1] and by J. Czekanowski.[2] Mollison has discussed particularly the problem of differences between two types, and he gives an arbitrary formula which later on was modified by St. Poniatowski,[3] according to whom the difference

$$D = (a_1 - a_2)\,\frac{\sigma_1 + \sigma_2}{\sigma_1\ \sigma_2}.$$

In this formula the distance between the two averages is expressed as a multiple of the two values of σ, and, as a measure of the distance, is assumed the value $\frac{\sigma_1 \sigma_2}{\sigma_1 + \sigma_2}$. The problem, however, can hardly be solved by such arbitrary methods.

In the following pages I shall discuss some possible approaches to the problem.

What we call difference in this case is not by any means an arithmetical difference; it is a judgment of the degree of dissimilarity of two series. If two series are so far apart that notwithstanding their variability they do not overlap, they are entirely dissimilar. If they do overlap they will be the more dissimilar, the less the amount of overlapping. In this sense we may say that two pairs of series in which the amount and character of overlapping are equal will be equally dissimilar. While we may thus determine equality of dissimilarity we are not in a position to determine quantitatively the degrees of dissimilarity.

In treating this problem we may first of all explain the meaning of similarity and dissimilarity by means of a few examples. Let us as-

[1] *Morphologisches Jahrbuch*, Vol. 42, p. 79.
[2] *Korrespondenz-Blatt der Deutschen anthropologischen Gesellschaft*, Vol. 40.
[3] *Archiv für Anthropologie*, N. F. Vol. X, p. 274.

sume that a pure negro and a pure white population are to be compared. The types are so distinct in all their features that in comparing them we should emphasize simply their dissimilarities. Now let us assume that a third community is added, consisting perhaps of baboons. It appears at once that our point of view would be shifted from a consideration of dissimilarities between negroes and whites to the similarities which they have in common as compared with the baboon, and their similarities will appear to us now under a new angle and as of different value.

When we compare a group of blond, blue-eyed North Europeans with dark complexioned, brown-eyed South Europeans, their dissimilarities are the most striking feature. If we add a negro community to these two groups, the similarities between the North and South Europeans would be much more prominently in our minds. We may observe the same changing attitude when we speak of family resemblances, or similarities. When we consider the children of a family, entirely by themselves, without any reference to any other family, they will appear to us as dissimilar. If the family has a particular characteristic feature, let us say, for instance, a long narrow nose, which all the children have to a greater or less extent, this will become the feature which makes them similar as compared to the rest of the population.

It is, therefore, clear that the concept of the degree of similarity depends upon the characteristics of all the groups that are under consideration and will change with the groups that are being compared.

In investigations on heredity it has been customary to determine the degree of similarity by means of the coefficient of correlation. When, for instance, parents and offspring are compared, the coefficient of correlation between the two will indicate the degree of their similarity. There is a biological relation between parent and offspring. The average form of the offspring is determined by the degree to which the parent differs from the average of the population to which he belongs. In marriage we may have selective mating through which the forms of two parents may be correlated. When the husband differs from the average of the population by a certain amount his wife will differ by a correlated amount. In both of these cases there is a functional relation between the two values. The distinguishing feature of fraternal correlation is that we are dealing with a natural group in which there is no true functional relation whatever. In a very large fraternity the bodily form of one member does not influence in any way either the average body form of the rest of the fraternity or the distribution of the individual forms. The fundamental equation [1] of correlation

$$[y] = qx.$$

[1] Brackets indicate averaging.

has no meaning in this case because the coefficient of regression q will approximate zero—the more so the larger the fraternity. This is due to the fact that the members of the fraternity are all members of the same variable class, while in all the other cases previously noted we are dealing with relations between different classes. Fraternal correlation originates only in a population in which the fraternities represent different types. If all the families had the same average value there would be no correlation and no similarity between brothers. The greater the heterogeneity of the population, the greater will be the correlation and similarity between members of a fraternity. If we call the coefficient of fraternal correlation r, the deviation of individuals from the average of the whole series x, the deviation of the family average from the general average of the population d, then for families of n children,

$$\left[\left(\frac{x_1+x_2+x_3+\ .\ .\ .\ +x_n}{n}\right)^2\right]=[d^2]. \tag{1}$$

If the standard deviation of all the values of d is s and the standard deviation of all the values of x is σ, then it follows from (1) that

$$\sigma^2\left(\frac{1+\overline{n-1}\,r}{n}\right)=s^2 \tag{2}$$

and for a very large number of n,

$$r=\frac{s^2}{\sigma^2}.$$

When all the families are alike, s is zero, and therefore r is also zero. The greater is s, that is the greater the diversity of family types, the larger is r.

If we assume the variabilities of the fraternities to be the same for all families and to be expressed by s_1, then for large numbers of n,

$$\sigma^2=s^2+s_1^2$$

and

$$r=\frac{1}{1+\frac{s_1^2}{s^2}}.$$

Exactly the same considerations may be made for racial types. A local variety may be considered as a fraternal group. The term r will then be a measure of the heterogeneity of the people or of the dissimilarity of the groups.[1]

[1] In another place ("On the Varieties of Lines of Descent Represented in a Population," *American Anthropologist*, NS. Vol. 18, pp. 1–9), I have shown how the heterogeneity of a population may be determined according to this method even in those cases in which each fraternity consists of a small number of individuals, and I have given a number of examples of heterogeneous and homogeneous

Thus we are led again to the conclusion that similarity has a meaning only when we consider a group as a whole, and that it is not possible to employ the terms "similarity" or "dissimilarity" in such a way that they have the same meaning for all combinations of groups. I have calculated, according to the formula just given, on the basis of Livi's data, the heterogeneity of the Italian population with regard to their cephalic index and find it to be 0.34, while it is only 0.12 for the Italian population north of the Appenines.

It does not seem that we can obtain in this manner a satisfactory measure for differences in the degree of similarity between single groups.

The problem of the definition of similarities has been treated fully in experimental psychology. Weber's law is actually an expression of the observation that the differences between two pairs of sensations are judged to be equal. In this case the basis of empirical determination of similarity is the probability of mistaking one difference for another. It is not, as was originally assumed, a measure of quantitative value of the sensation itself. This concept of similarity holds good not only in the case of simple sensations but also in the field of more complex experience. We may speak of similarity, or of the probability of failing to differentiate, for the most diverse kinds and the most complex forms of mental experience. The problem that we are discussing here has suggested itself in every comparative study of mental processes.

In an analogous manner we may define the degree of similarity as the probability of mistaking an individual that belongs to one group for a member of any of the other groups concerned. The degree of dissimilarity may then be determined by the probability of recognizing an individual as belonging to his own group.

If we call p_1, the probability of occurrence of a certain measurable

---

populations. It may be remarked here that in a mixed population for which no information as to descent is available, heterogeneity will ordinarily appear when the measurements of two organs are correlated. If various families represent different types with regard to two measurements which are not correlated; if the averages of these types have deviations d and d' from the general average, and if the individuals deviate from the respective averages of their families by the amounts x and y, then their correlation may be expressed by

$$r = \frac{[(d+x)\,(d'+y)]}{\sigma x \cdot \sigma y},$$

and since x and y are assumed not to be correlated

$$r = \frac{dd'}{\sigma x \cdot \sigma y}.$$

Calculating in this way the correlation between cephalic index and stature for the whole of Italy on the basis of Livi's data, we find a correlation of +.1, while in a homogeneous series the correlation between stature and cephalic index is negative. The value +.1 is derived from observations of 300,000 individuals, and is therefore quite certain.

I have discussed this phenomenon in relation to the population of Paris in the *American Anthropologist*, NS. Vol. 1, pp. 448-61.

form X, which is a member of a certain variable class; p_2 the probability of occurrence of the same form X as a member of a second class, etc.; if, furthermore, n such classes occur in which the observation X has the probabilities of occurrence $p_1, p_2, p_3 \ . \ . \ . \ p_n$ and if the probabilities of occurrence of the classes are $p_1', p_2', p_3' \ . \ . \ . \ p_n'$, then the probability of occurrence of the form X as a member of the series r is $p_r p_r'$; and the total probability that any individual of measure X will occur in the aggregate of n series will be

$$p_1 p_1' + p_2 p_2' + p_3 p_3' + \ . \ . \ . \ + p_n p_n'.$$

The probability that a certain value X among all the values of X belongs to the series r is therefore

$$\frac{p_r p_r'}{\Sigma p p'}.$$

Now let us assume that there be given n individuals of the form X whose provenience is not known, but who actually belong to the series r. If our knowledge of the aggregate is perfect, then the n given individuals must be assigned according to the principle that among n assignments $\frac{p_r p_r'}{\Sigma p p'} n$ must belong to the series r, the rest to the series *not* r. The probability of a correct assignment of a single observation is therefore

$$\frac{p_r p_r'}{\Sigma p p'}.$$

If, now, a series of observations of the class r is given, the probability of the occurrence of X in this series will be p_r.

Then, under the conditions just mentioned, the probability of the occurrence and at the same time of the correct assignment of X to the class r will be

$$P_{x,r} = \frac{p_r^2 \, p_r'}{\Sigma p \ p'}. \tag{3}$$

The probability that any one of all the values of X that do occur is correctly assigned, is therefore

$$P_r = \int_{-\infty}^{+\infty} \frac{p_r^2 \, p_r'}{\Sigma p \ p'} dx, \tag{4}$$

and we may call this the degree of dissimilarity between the series r and the aggregate of all the series that are being considered. It is obvious that in this manner the degrees of diversity for all the other individual series may be established, that they may be subtracted from

one another, and that in this manner differences in the degree of similarity may be determined.

It may be pointed out that when three series are compared in this manner in pairs, the resultant values are not additive. If only series (1) and series (2), then series (1) and (3), then series (2) and (3) are considered, the sum of the difference between (1) and (2) plus that between (2) and (3) will not be equal to the difference between (1) and (3). This is another expression of the observation made before that the meaning of similarity changes with the aggregate of the series that is being considered.

It might also seem possible to arrange the single series in the order of their averages and to determine their dissimilarities step by step. Here the difficulty may arise that two succeeding averages may be nearly the same, while their variabilities may be quite different. Whenever this occurs quite an erroneous impression of the differences would be given. The reason for this difficulty lies in the fact that the difference as here defined depends upon the averages and variabilities of the single series, and that certain combinations of these result in the same degree of dissimilarity.

In the case treated here the various series enter into the aggregate according to the number of individuals representing each series. It might be, for instance, that a large mass of material has been accumulated for one group and that another group is known through the study of a very few individuals only. Our expression contains, therefore, a weighting according to number which obscures the more general theoretical question. If the groups were known perfectly, then all would have equal weight, *i. e.*, we should have to assume them to be represented by equal numbers. All the values p' would then be equal. This is the assumed condition under which most comparisons of types are actually made. We have, then,

$$P = \int_{-\infty}^{+\infty} \frac{p_r^2}{\Sigma p} dx \tag{5}$$

as a measure of dissimilarity and

$$P_1 = 1 - \int_{-\infty}^{+\infty} \frac{p_r^2}{\Sigma p} dx \tag{5*}$$

as a measure of similarity.

Whether equation (5) or equation (4) represents the conditions of a set of observations, or whether a form with still different weighting should be selected, depends upon the clarity of our concept of the characteristics of each group. If we assume each group as thoroughly

studied and therefore known in all its characteristics, then equation (5) will represent the conditions adequately. On the other hand, if we are impressed by the unclassified series as a whole, without detailed study of each group, and if we try to determine the similarities and dissimilarities on this basis, then formula (4) will correspond to the conditions of the investigation. If subjective elements are to be eliminated as far as possible, we must try to adjust conditions so that equation (5) can be applied. As a matter of fact, our judgment of similarity in all cases of this type is fluctuating; sometimes one group, sometimes another, is most prominently in our minds, and the actual assignments are therefore different from the two extreme forms (4) and (5) discussed before and may lie somewhere in between, or they may change with changing mental conditions. The more thorough our knowledge of each series, the closer will be the approach to form (5).

The formulas here given present the inconvenience that the values obtained for similarity are the smaller, the larger the number of series forming the aggregate, so that when the number of similar series is very great the values of their similarities will be exceedingly small.

In the final results it may appear that some of these series have the same degree of dissimilarity. If the averages and variabilities of these series are also indicative of identity, the series should be combined.

It must be remembered that it is possible for a number of different distributions to result in the same amount of dissimilarity. Since every distribution depends at least upon two constants, average and standard deviation, there are whole sets of functions which will give us the same value for the total probability of mistaking a member of one series for a member of the rest of the aggregate. However, owing to the general likeness of forms of distribution, the occurrence of this event is improbable. On the other hand, dissimilarity can occur only when distributions are unlike. The minimum amount of dissimilarity is found when all the series are identical. If there are n series, the value of dissimilarity will be $\frac{1}{n}$.

All the dissimilarities have necessarily positive values. For this reason if one dissimilarity consists in the prevalence of high values, another one in the prevalence of low values, they ought to be distinguished, and the direction of dissimilarity may be indicated by opposite signs. It will, therefore, be necessary in every case to consult also the characteristic averages and distributions in order to give the proper sign to the values of dissimilarity. It must, however, be understood that the values are additive only without regard to the sign, since they all represent probabilities.

I have calculated on the basis of (5) the dissimilarities of the cephalic-index, stature, and hair color, in Italy, basing my calculations on Livi's material.[1] The following tables give the results of this calculation:

TABLE I, CEPHALIC INDEX

	Dissimilarity	Average cephalic index	Average cephalic index of individuals judged to belong to each series [a]
Piemont	0.103	85.9	87.8
Liguria	(−) 0.061	82.3	82.3
Lombardy	0.083	84.4	85.2
Venice	0.088	85.0	86.2
Emilia	0.091	85.2	86.8
Tuscany	(−) 0.063	82.3	82.4
Marche	0.073	84.0	85.4
Umbria	0.074	84.1	85.5
Lazio	(−) 0.067	81.0	80.0
Abruzzi	(−) 0.065	81.9	81.7
Campania	(−) 0.068	82.1	82.0
Puglie	(−) 0.080	79.8	78.4
Basilicata	(−) 0.072	80.8	79.9
Calabria	(−) 0.110	78.4	75.4
Sicily	(−) 0.085	79.6	78.4
Sardinia	(−) 0.152	77.5	74.9

[a] See p. 438.

TABLE II

	Stature			Hair Color
	Dissimilarity	Average stature	Average stature of individuals judged to belong to each series	Dissimilarity
Piemont	0.0643	164.9	165.7	0.070
Liguria	0.0672	165.5	166.8	0.069
Lombardy	0.0667	165.3	166.6	0.066
Venice	0.0751	166.6	168.6	0.072
Emilia	0.0673	165.3	166.8	0.066
Tuscany	0.0682	165.6	167.1	0.070
Marche	(−) 0.0633	163.8	163.5	0.065
Umbria	(−) 0.0636	164.2	164.4	0.064
Lazio	(−) 0.0630	164.3	164.4	(−) 0.068
Abruzzi	(−) 0.0654	163.2	162.3	0.068
Campania	(−) 0.0643	163.5	163.0	(−) 0.065
Puglie	(−) 0.0646	163.5	163.0	(−) 0.064
Basilicata	(−) 0.0689	162.6	161.3	(−) 0.068
Calabria	(−) 0.0652	163.1	162.3	(−) 0.066
Sicily	(−) 0.0636	163.5	162.9	(−) 0.065
Sardinia	(−) 0.0754	161.9	160.4	(−) 0.077

I have calculated these tables from the summary of observations for sixteen provinces. The minimum of dissimilarity is therefore 0.06275, and it will be seen that a number of provinces differ so little from the

[1] Rodolfo Livi, *Antropometria Militare*, pp. 248–49; 266–67; 286–87.

mass of the Italian material that they have no distinguishing characteristics. On the other hand, Piemont and Sardinia show, comparatively, very great differences.

If a number of independent traits are examined together the probability of a correct assignment increases. When the probability of correct assignments for various traits are $p_1, p_2 \ . \ . \ . \ p_r$, then the probability of an incorrect assignment considering all the characteristics together will be

$$(1-p_1)(1-p_2) \ . \ . \ . \ (1-p_r).$$

When all the series are identical, and only one trait is studied, the probability of an incorrect judgment will be $\frac{n-1}{n}$, where n is the number of series compared. For r independent traits, the probability of an incorrect judgment in identical series will be $\left(\frac{n-1}{n}\right)^r$. This value represents the maximum probability of an incorrect judgment. The minimum probability of an incorrect judgment occurs in cases of absolute difference of all individuals, and is zero.

The values of probabilities derived from a combination of several characteristics may be made more comparable by determining the ratio between $\left(\frac{n-1}{n}\right)^r$ and the observed value of the product $(1-p_1)(1-p_2) \ . \ . \ . \ (1-p_r)$.

Applying this to the sixteen provinces of Italy, tabulated on page 435, we find the maximum probability of misjudgment for a single characteristic to be $\left(\frac{15}{16}\right)$, or 0.937; for two characteristics combined, $\left(\frac{15}{16}\right)^2$, or 0.879; for three characteristics combined, $\left(\frac{15}{16}\right)^3$, or 0.824.

Table III gives the absolute probabilities of misjudgments (column I) and their ratio to the maximum probability of a misjudgment based on a single character, namely, 0.937 (column II in the first three double columns). Column II in the fourth double column gives the ratio between the maximum probability of a misjudgment based on a comparison of two characters and the observed probability of a misjudgment for head index and stature. Column II in the last double column gives the same for three characters, namely, head index, stature, and hair color.

In this series Piemont and Sardinia appear as the most distinct areas. It seems then justifiable to compare these alone in order to determine their dissimilarity. According to the same method, applied

TABLE III

	Cephalic index		Stature		Hair color		Cephalic index and stature		Cephalic index, stature, and hair	
	I	II	I	II	I	II	I	II	I	II
Piemont	0.897	0.957	0.936	0.999	0.930	0.993	0.840	0.956	0.781	0.948
Liguria	0.939	1.000	0.933	0.996	0.931	0.994	0.876	0.997	0.816	0.990
Lombardy	0.917	0.979	0.933	0.996	0.934	0.997	0.855	0.973	0.799	0.970
Venice	0.912	0.973	0.925	0.987	0.928	0.990	0.844	0.960	0.783	0.950
Emilia	0.909	0.970	0.933	0.996	0.934	0.997	0.848	0.965	0.792	0.961
Tuscany	0.937	1.000	0.932	0.995	0.930	0.993	0.873	0.993	0.812	0.985
Marche	0.927	0.989	0.937	1.000	0.935	0.998	0.869	0.989	0.813	0.987
Umbria	0.926	0.988	0.936	0.999	0.936	0.999	0.867	0.986	0.811	0.984
Lazio	0.933	0.996	0.937	1.000	0.932	0.995	0.874	0.994	0.815	0.989
Abruzzi	0.935	0.998	0.935	0.998	0.932	0.995	0.874	0.994	0.815	0.989
Campania	0.932	0.995	0.936	0.999	0.935	0.998	0.872	0.992	0.815	0.989
Puglie	0.920	0.982	0.935	0.998	0.936	0.999	0.860	0.978	0.805	0.977
Basilicata	0.928	0.990	0.931	0.994	0.932	0.995	0.864	0.983	0.805	0.977
Calabria	0.890	0.950	0.935	0.998	0.934	0.997	0.832	0.947	0.777	0.943
Sicily	0.915	0.977	0.936	0.999	0.935	0.998	0.856	0.974	0.800	0.971
Sardinia	0.848	0.905	0.925	0.987	0.923	0.985	0.784	0.892	0.724	0.879

to these two territorial divisions alone, we find the dissimilarities as shown in Table IV.

Thus, the probability of a misjudgment for Sardinians and Piemontese for cephalic index alone is 0.420 of the probability for identical series; for cephalic index and stature it is 0.389; and for cephalic index, stature, and hair color it is 0.342 of the probability for identical series. Still more strictly speaking, we may express the last column in the following way: In attempting a classification of individuals belonging to two identical series for which three measurements are compared, $(\frac{1}{2})^3$ out of the total number, or 125 out of 1,000 individuals, will be misplaced, but only 43 out of 1,000 are misplaced when comparing Sardinians and Piemontese.

The method here outlined enables us to segregate the most distinctive groups from the mass of overlapping material that belongs to a large area, as we have segregated here the Piemontese and Sardinians.

TABLE IV

	Cephalic index		Cephalic index and stature		Cephalic index, stature, and hair color	
	I	II	I	II	I	II
Cephalic index	0.210	0.420	....	....	....	....
Stature	0.463	0.926	0.097	0.389	0.043	0.342
Hair	0.438	0.976	....	....	....	....

It will be noted that in our example Venice and Calabria are also strongly divergent in type. A study of the measurements of these provinces shows that Calabria may be combined with Sardinia, while Venice appears somewhat distinct from Piemont. Accordingly, Calabria and Sardinia should be united. By a gradual elimination of intermediate forms it is possible not only to segregate the most divergent types of an area but also to determine the degree of their similarity.

In applying the fundamental thought underlying our considerations to the classification of mankind, we might ask ourselves which are the series for which the similarity or the probability of a misjudgment becomes zero, and these might be considered as the present fundamental human types. A satisfactory solution of this problem must not be based on the consideration of a few standardized measurements, but the features to be studied must be selected after a careful investigation of what is most characteristic of each group.

It is also feasible to find in this manner outstanding types of a definite area and to arrange them according to the degrees of their similarity. The interpretation of the similarity, whether due to mixture, environment, or other causes, is of course a purely biological problem for which the statistical inquiry furnishes the material but which cannot be solved by statistical methods.

We have seen that the individuals of a definite bodily form are not all assigned by us to the group to which they belong. The impression which we receive of characteristic forms of a particular series depends upon the distribution and the forms of individuals which we assign to it. While their distribution does not conform to the exponential function, it does not differ fundamentally from the distribution of other variable phenomena, and for this reason our impression of the general characteristic form of the series is expressed by the average of individuals that we assign to it. This value is

$$a = \int_{-\infty}^{+\infty} \frac{X p_r^2 p_r'}{\Sigma p'} dx, \tag{6}$$

or, if the values of p^1 are equal,

$$a = \int_{-\infty}^{+\infty} \frac{X p_r^2}{\Sigma p} dx. \tag{6*}$$

This calculation for Italy has been carried out on the basis of the previous tables. The psychological impression of characteristic head forms of each province of Italy is given in the last column of the table on page 435, while the second columns contain the actual averages.

This table is suggestive because it shows that we receive an exaggerated impression of the characteristics of a series, because individuals that are similar to other series are assigned to them according to their appearance and are merged in the general background represented by the aggregate. Our impression, however, does not correspond to an actual type. This proves that the attempts to analyze a series according to similarities of individuals into a number of subtypes is methodologically not admissible, and that all subdivisions must be based on the study of the series as a whole, not upon selected types.

The chief difficulty in the practical application of the method outlined in the preceding pages is due to the facts that the degree of similarity depends upon the aggregate treated, and that there is no relation between the numerical values obtained for different aggregates. Not even the equality of differences between several given series need persist if new members are added to the aggregate or are taken away from it.

In cases of continuous changes of a type from one extreme form to another, an artificial classification of the aggregate is unavoidable. By means of repeated adjustment equal degrees of similarity might be found according to the method outlined here, but the actual carrying out of such a plan offers serious difficulties. In such cases each series might be considered as a specialized form of the general aggregate and compared with it. The aggregate itself may, however, be established in two different ways. We may disregard the number of existing individuals, considering each morphological type contained in the aggregate as a unit. The units would then be given equal weight (*i. e.*, equal numbers of cases). Or we may take the whole series as it exists at the present time, counting the total number of individuals that it contains, regardless of local types that may represent the same morphological form. By either of these methods we ascertain how dissimilar each morphological type is from the aggregate, but these values cannot be used to determine the mutual dissimilarities of the single series contained in the aggregate. When the types are combined according to the present actual number of individuals representing them, the most numerous type will appear least distinct from the average, merely on account of the large number of its members. This difficulty can hardly be avoided by comparing each series with the aggregate of the remaining series, because by this method the standard of comparison is always changing. On the other hand, the formation of the aggregate by giving equal weight to each morphological type entails the difficulty that we tried to avoid, namely, an arbitrary classification of the groups as a number of morphological types.

The problem may be approached in another manner. We may determine the frequency distribution of the differences between individuals belonging to one series and those belonging to all the other series of the aggregate. In this inquiry we have to determine the average difference between the representatives of one series and those of all the other series, and the variability of this difference. When the series are arranged in pairs, the differences between the averages are additive, but the variabilities are not comparable. The difference between a series (1) and a series (2) will have the variability $\pm\sqrt{\sigma_1^2+\sigma_2^2}$, so that if one series is taken as the basis of comparison with all the others, the variabilities have a significance only in relation to the one series which has been arbitrarily selected as a standard.

The interrelations between the series can be determined only when we consider any one series in relation to all the others. Assuming again that all the series are known, we may call the values of the particular series

$$\begin{array}{c} a_1+x_1 \\ a_2+x_2 \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ a_n+x_n. \end{array}$$

The differences between one series and all the others taken successively will be

$$\begin{array}{c} (a_1-a_2)+(x_1-x_2) \\ (a_1-a_3)+(x_1-x_3) \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ (a_1-a_n)+(x_1-x_n), \end{array}$$

or since there are $n-1$ groups of this type, the average difference will be

$$\frac{n}{n-1}a_1-\frac{\Sigma a}{n-1}.$$

The mean square variability of this difference will be

$$\frac{1}{n-1}\left\{(x_1-x_2)^2+(x_1-x_3)^2+ \quad . \quad . \quad . \quad +(x_1-x_n)^2\right\},$$

or

$$\frac{n-2}{n-1}\sigma_1^2+\frac{\Sigma\sigma^2}{n-1}.$$

We have, therefore, the characteristic difference between series (1) and the rest of the series

$$\frac{n}{n-1}a_1-\frac{\Sigma a}{n-1} \quad \sqrt{\frac{n-2}{n-1}\sigma_1^2+\frac{\Sigma\sigma^2}{n-1}}.$$

If we assume a distribution according to the probability function, the differences may be made comparable by measuring them by their standard deviations. The difference between series (1) and all the other series taken conjointly will therefore be

$$\frac{na_1-\Sigma a}{\sqrt{(n-1)(n-2\sigma_1^2+\Sigma\sigma^2)}}, \tag{7}$$

or

$$\frac{n}{n-1}\,\frac{a_1-[a]}{\sqrt{[\sigma^2]+\sigma_1^2+\dfrac{[\sigma^2]-\sigma_1^2}{n-1}}}.$$

If n is sufficiently large, we may write approximately

$$\frac{a_1-[a]}{\sqrt{[\sigma^2]+\sigma_1^2}}, \tag{8}$$

which is the difference between the average of series (1) and the average of the total aggregate including series (1) measured by its variability. From this value and by means of a table of the probability integral may be determined the probability that the difference between an individual belonging to series (1) and any other individuals of the aggregate has any selected value, particularly the probability that this difference may be greater or less than zero.

The advantages of this definition of difference is the relative stability of the values obtained when the series are arranged in different groups. As an instance I give the calculation for Italy. Table V gives the averages and variabilities of the cephalic index for each circondario. In each column are given the number of occurrences of the values of the average cephalic index by half units, as 76.0–76.4, 76.5–76.9, etc.

TABLE V

Variability	Cephalic index																			
	76		77		78		79		80		81		82		83		84		85		86		87	
3.00–3.24	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	3	1	.	.	.	.	.	.
3.25–3.49	.	1	.	.	1	2	4	1	2	1	.	1	.	4	2	1	3	3	5	2	1	1	.	.
3.50–3.74	.	.	2	4	6	1	2	6	6	4	2	6	3	4	4	1	6	4	6	3	4	2	.	2
3.75–3.99	.	.	2	1	1	2	1	3	.	3	6	4	1	2	5	3	1	3	6	3	3	4	.	.
4.00–4.24	.	.	.	1	1	1	2	1	2	1	.	3	1	2	2	1	2	1	1	1	.	2	.	.
4.25–4.49	.	.	.	.	.	.	.	.	.	.	.	2	.	.	1	1	2	1	.	1	.	.	.	.
4.50–4.74	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.

Considering each combination of an index and a variability occurring in Table V as a single morphological group, we obtain the values given in Table VI for formula (8) for 84 such groups. Considering

TABLE VI

Variability	Cephalic index											
	76	77	78	79	80	81	82	83	84	85	86	87
3.00–3.24...	....	....	....	....	.... −.33	....	....	....	+.38 +.49	....	....	....
3.25–3.49...	−1.21	....	−.82 −.72	−.62 −.52	−.42 −.32	.... −.12	−.02	+.18 +.28	+.37 +.47	+.57 +.67	+.77 +.87	....
3.50–3.74...	....	−.98 −.88	−.79 −.69	−.60 −.50	−.40	−.21 −.12	−.02 +.07	+.17 +.27	+.36 +.46	+.55 +.65	+.74 +.84	....+1.03
3.75–3.99...	....	−.94 −.85	−.76 −.67	−.57 −.48	−.39	−.20 −.11	−.02 +.07	+.16 +.26	+.35 +.44	+.53 +.62	+.72 +.81	....
4.00–4.24...	....	.... −.82	−.73 −.64	−.56 −.47	−.38 −.29	.... −.11	−.02 +.07	+.16 +.25	+.34 +.43	+.52 +.61	+.69	....
4.25–4.49...	....	....	....	....	....	.... −.11	....	+.15 +.24	+.33 +.41	+.50 +.59	....	....
4.50–4.74...	....	....	....	....	....	....	....	.... +.23	....	....	....	....

TABLE VII

Variability	Cephalic index											
	76	77	78	79	80	81	82	83	84	85	86	87
3.00–3.24...	....	....	....	....	.... −.36	....	....	....	+.36 +.46	....	....	....
3.25–3.49...	−1.24	....	−.84 −.74	−.65 −.55	−.45 −.35	.... −.15	−.05	+.15 +.25	+.35 +.45	+.55 +.65	+.75 +.84	....
3.50–3.74...	....	−1.01 −.91	−.82 −.72	−.63 −.53	−.43 −.34	−.24 −.14	−.05 +.05	+.14 +.24	+.34 +.43	+.53 +.63	+.72 +.82	.... +1.01
3.75–3.99...	....	− .98 −.88	−.79 −.70	−.60 −.51	−.42	−.23 −.14	−.05 +.05	+.14 +.23	+.32 +.42	+.51 +.60	+.70 +.79	....
4.00–4.24...	....	.... −.85	−.76 −.67	−.58 −.49	−.40 −.31	.... .13	.04 +.04	+.13 +.22	+.31 +.40	+.49 +.58	+.67	....
4.25–4.49...	....	....	....	....	....	.... .13	....	+.13 +.22	+.30 +.39	+.48 +.57	....	....
4.50–4.74...	....	....	....	....	....	....	....	.... +.21	....	....	....	....

TABLE VIII

Variability	Cephalic index											
	76	77	78	79	80	81	82	83	84	85	86	87
3.00–3.49	−1.14	....	−.74	−.54	−.34	−.14	+.06	+.26	+.46	+.66	+.86	....
3.50–3.99	....	−.88	−.69	−.51	−.32	−.13	+.06	+.24	+.43	+.62	+.81	+.99
4.00–4.49	....	−.83	−.65	−.48	−.30	−.12	+.05	+.23	+.40	+.58	+.76	....
4.50–4.99	....	....	....	....	....	....	....	+.21	....	....	....	....

TABLE IX

Variability	Cephalic index											
	76	77	78	79	80	81	82	83	84	85	86	87
3.00–3.49	−1.21	....	−.81	−.60	−.40	−.20	±0.00	+.20	+.40	+.60	+.81	....
3.50–3.99	....	−.94	−.75	+.57	−.38	−.19	± .00	+.19	+.38	+.57	+.75	+ 94
4.00–4.49	....	−.88	−.71	−.53	−.35	−.18	± .00	+.18	+.35	+.53	+.71	....
4.50–4.99	....	....	....	....	....	....	....	+.17	....	....	....	....

the population in its actual distribution, *i. e.*, counting each morphological group according to the actual number of its occurrence, we get for 206 groups the figures given in Table VII. For whole units of cephalic index and half units of variability the corresponding figures for 32 morphological units are given in Table VIII; and counting the actual frequencies of the single groups we obtain the figures given in Table IX.

The problems take a slightly different form when populations are compared with regard to features that occur in a certain percentage of individuals and are absent in the rest. If, for instance, one population consists of 15 per cent negroes and 85 per cent whites, another one of 30 per cent negroes and 70 per cent whites, it might seem that the difference could be stated simply as a difference of 15 per cent, but obviously the dissimilarity of these two types of population would not be the same as in another pair in which we have 40 per cent negroes and 60 per cent whites in one and 55 per cent negroes and 45 per cent whites in the other. In the latter case the populations would seem more alike to us than in the former case. This difficulty is still more pronounced if there are present not merely two types but a larger number of types in varying proportions. In all these cases we may apply the same methods which we used for the determination of similarity of measurable quantities.

The first method of determining similarities may be applied to this problem without change. If a certain feature has the probabilities $p_1, p_2, p_3 \ldots p_n$ in a number of series, then the frequency of its occurrence in any one of the n series will be $p_1+p_2+p_3+ \ldots +p_n$. The probability that a single individual of series r whose provenience is not known is assigned correctly will be, therefore, $\frac{p_r}{\Sigma p}$, and if the probability of occurrence of an individual of this series is pp', then the total probability of a correct assignment will be $\frac{p_r^2 p_r'}{\Sigma pp'}$, or if we disregard the number of cases in each series, $\frac{p_r^2}{\Sigma p}$. The sum of these values for all the distinct traits occurring in the population, that is, $\Sigma\frac{p_r^2 p_r'}{\Sigma p'p}$, will be the total probability of a correct assignment. In a case in which only two forms constitute a population, and comparing two populations only, we have accordingly for equal numbers of p'

$$\frac{p_1^2}{p_1+p_2}+\frac{(1-p_1)^2}{1-p_1+1-p_2}.$$

According to the second method we have to determine the average difference between individuals of the several series and its variability. If in a series the probability of the occurrence of a phenomenon is p, then those individuals in which it does occur, and who therefore have the probability of occurrence 1, have a deviation from the probability of the series $1-p$; those in whom it does not occur and who therefore have the probability of occurrence 0, have a deviation from the probability of the series $-p$. Of the former there are p cases, of the latter $1-p$ cases. We have, therefore, the following combinations:

The phenomenon is		Frequency	
present in series (1), present in (2)	p_1p_2	times	0
present in series (1), absent in (2)	$p_1(1-p_2)$	"	+1
absent in series (1), present in (2)	$(1-p_1)p_2$	"	−1
absent in series (1), absent in (2)	$(1-p_1)(1-p_2)$	"	0

The last column represents the average differences between the probabilities of occurrence for each combination.

The average is therefore

$$p_1(1-p_2)-p_2(1-p_1)=p_1-p_2.$$

We have, therefore, the following deviations from the averages and their squares:

Frequencies	Difference of each series from average (p_1-p_2)	Squares of differences
p_1p_2 times	$-(p_1-p_2)$	$(p_1-p_2)^2$
$(1-p_1)(1-p_2)$ "	$-(p_1-p_2)$	$(p_1-p_2)^2$
$p_1(1-p_2)$ "	$1-(p_1-p_2)$	$(1-p_1-p_2)^2$
$p_2(1-p_1)$ "	$-1-(p_1-p_2)$	$(1+p_1-p_2)^2$

and by addition

$$\sigma^2=p_1(1-p_1)+p_2(1-p_2).$$

When we compare in this manner p_1 with all the other series, we obtain as before

$$a=\frac{n}{n-1}p_1-\frac{\Sigma p}{n-1}$$

$$=\frac{n}{n-1}(p_1 1-[p])$$

$$\sigma=\pm\sqrt{\frac{n-2}{n-1}p_1(1-p_1)+\frac{\Sigma p(1-p)}{n-1}}$$

$$=\pm\sqrt{p_1(1-p_1)+[p(1-p)]+\frac{[p(1-p)]-p_1(1-p_1)}{n-1}}.$$

Since $\sqrt{p(1-p)}$ is the standard deviation for p, the formula appears identical with the one previously derived.

THE LONG-TIME CYCLES OF PANDEMIC INFLUENZA[1]

By Otto R. Eichel, *New York State Department of Health*

The object of this paper is to present briefly some observations on the occurrence, in point of time, of the great influenza pandemics, and to indicate certain phases, the statistical study of which may prove valuable. For this purpose the term "pandemic" is applied to those great epidemics which have not only affected a whole country, continent, or hemisphere, but have, in a series of great waves, encircled the entire globe. The term "influenza" is here used to designate the disease or clinical entity usually known by that name to physicians; it refers also to the same disease known by other names before the term influenza came into general use.

The study of these phenomena in the light of history is of course limited in precision by the absence of mortality statistics in modern form previous to about the last century and a half. Nevertheless, the pandemics stand out so clearly in the history of the last four centuries that one may endeavor to measure them roughly and to interpret their relationship to each other. Viewed in this perspective they present a picture characterized by unmistakable features.

Our sources of data are chiefly the descriptive and critical studies of Hirsch, Creighton, Leichtenstern, Thompson, and others, and official vital statistics. Previous to the seventeenth century the records were made for the most part by lay chroniclers; subsequent to the days of the great Sydenham, in the early seventeenth century, physicians in increasing numbers recorded their findings; and during the nineteenth century there accumulated an extensive literature including descriptive official statistics. It must be noted that before the last century the scientific writers on influenza, chiefly physicians, compensated for their lack of the precise methods and resources of modern clinics, laboratories, and official records, not only by comprehensive and minute observations, but also by a thorough mastery of language in presenting them.

The method used to identify and relate to each other the historic influenza epidemics cannot be stated in exact detail in this brief paper, but it will be described sufficiently to make clear the procedure. We selected the records of England and Wales as a measure of the pande-

[1] Read at the Eighty-third Annual Meeting of the American Statistical Association at Pittsburgh, Pa., December, 1921.

micity of influenza because they cover a long period of time and are readily available, including critical studies of them, especially by Creighton.

The first step is to chart the annual influenza death rates of England and Wales since 1838 from data in the yearly reports of the Registrar-General, with the exception of a negligible interval of four years during which no epidemic occurred and for which figures are not available. The English "bills of mortality" were printed first in 1629, although the parish registers of deaths had been kept since 1538 and various parish bills of mortality were compiled as early as 1532. Next, from Creighton's study of the records and other sources of information, we are able roughly to estimate the mortality [1] of these various epidemics from 1510 to 1838: (1) of definite influenza; (2) of indefinite but very probable influenza; (3) of mild and doubtful character which may possibly have been influenza. These three classes must be held clearly in mind in our subsequent discussion. (See Charts I and II.)

In this connection, the nomenclature of influenza is important because throughout its history it appears under a great variety of names, many of them popular designations indicating either the severity or the mildness of the disease. In 1782 the term "influenza" was formally adopted by the College of Physicians in England and since then has been in general use in the English literature. But the wealth of medical and lay descriptions of the epidemics previous to 1782 establishes clearly the clinical character of the important influenza outbreaks.

For charting and comparison the epidemics may then be classified as follows, with regard to their relative severity and probable annual mortality from influenza per 100,000 population for the entire year in which each epidemic occurred:

1. Very mild; rate under 15.
2. Mild; rate over 15 and under 30.
3. Moderately severe; rate over 30 and under 45.
4. Very severe; rate 45 or over.

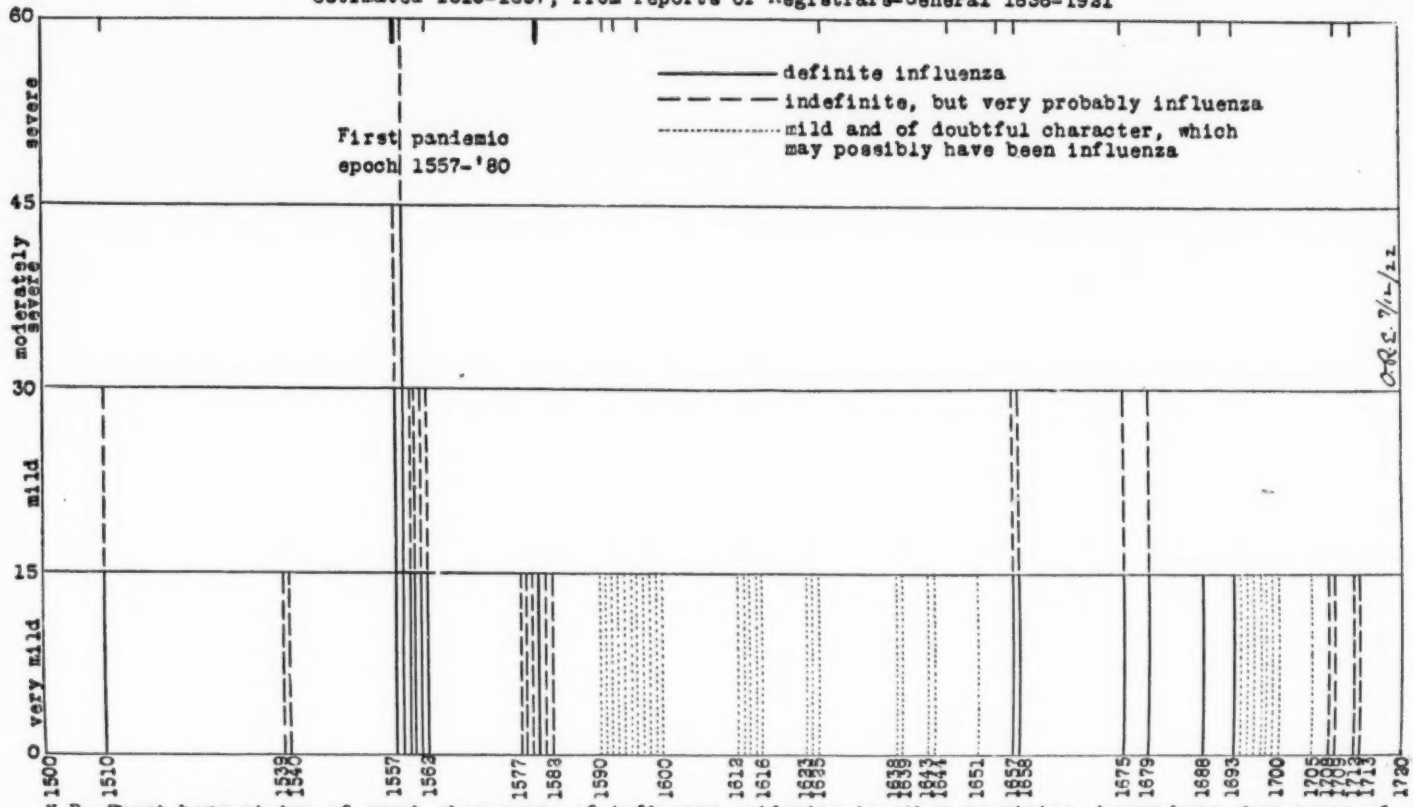
Although the definitions involve assumptions, derived from study of the annual statistics of the Registrars-General since 1838, in order to designate the relative severity of each epidemic preceding that year, it is not a matter of moment for our purpose whether such designations do accurately represent the real mortality. These groups might be reduced to 2 or 3, or extended to 6, and the corresponding mortality values likewise raised or lowered, and it would still be possible readily to allocate to them any of the historic epidemics with regard to their severity *as compared with each other* and with those which followed

[1] The term "mortality" is used in this paper to indicate deaths in proportion to population.

Chart I

THE LONG TIME CYCLES OF INFLUENZA IN ENGLAND AND WALES

Annual death rates per 100,000 population
estimated 1510-1837; from reports of Registrars-General 1838-1931



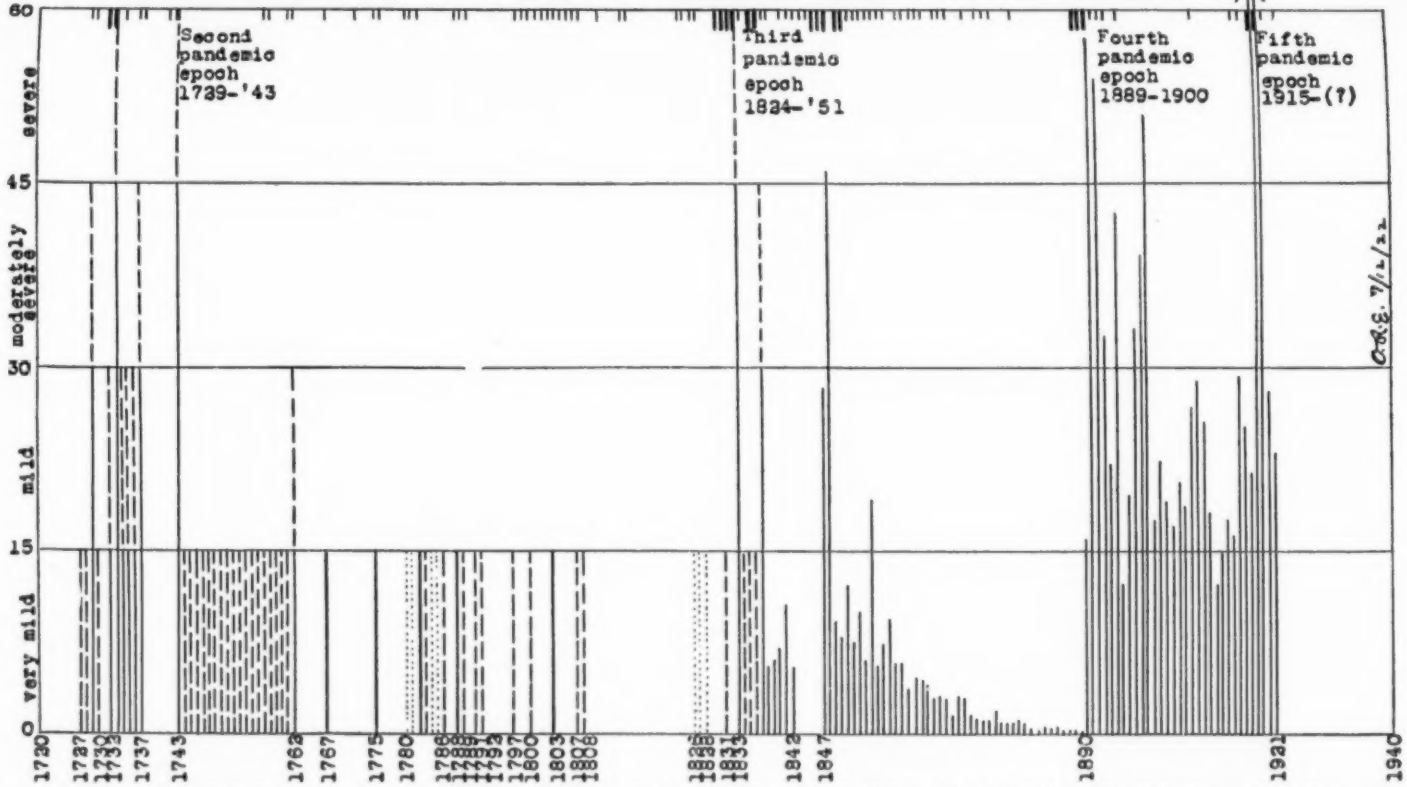
N.B. Short bars at top of graph show years of influenza epidemics in other countries; heavy bars show years of world-wide outbreaks, or real pandemics. (Also see foot note to Chart II.)

Chart II

THE LONG TIME CYCLES OF INFLUENZA IN ENGLAND AND WALES (Concluded from Chart I)

Chart II

THE LONG TIME CYCLES OF INFLUENZA IN ENGLAND AND WALES (Concluded from Chart I)



N.B. The broken parts of the continuous bars show the range within which the actual mortality probably fell, in both Charts I and II. (Also see foot note to Chart I.)

The actual rate for 1918 is over five times as high as that shown on the chart, and the rate for 1919 is nearly twice that represented.

1838. Even if opinion should differ as to whether a certain epidemic was mild or moderately severe, or otherwise, or whether or not it was true influenza, it would be found that it could be allocated approximately, according to the definitions formulated, both as to severity and to the certainty of its being true influenza. Allowing for wide differences of judgment, the resulting graphic pictures in each instance would still present the same significant features.

The data thus prepared are then plotted in their proper places on the chart constructed to show the annual influenza death rates since 1838 and according to the same scale.

Inasmuch as very extensive outbreaks [1] occurred which left little or no impress on the bills of mortality or the later official statistics, these questions now arise: (1) For the present purpose, is the mortality a reliable index of epidemicity? (2) Are great epidemics causing heavy mortality *in Britain* a reliable index to the pandemics?

The answers to these questions are evident from the fact that the years of severe outbreaks with high influenza mortality in Britain coincide with those periods during which the disease was encircling the globe and causing high mortality everywhere. Such pandemic periods or "epochs" (a term suggested by Leichtenstern) were followed by *occasional* less fatal outbreaks in the long interval of time before the next pandemic epoch began, some of which were very extensive, indicating a high degree of communicability, and affected large proportions of the people of a country, or even a continent or hemisphere.[2] Not only was the mortality low but the fatality (*i. e.*, percentage of fatal cases) also must have been low—a condition characteristic of the terminal phase of severe epidemics of various diseases. These mild and extensive epidemics did not coincide with periods of truly world-wide diffusion, but they are very confusing in the chronological history of the disease and seem to have obscured the picture of the real pandemic waves and their rise and subsidence. These waves form a remarkable and consistent sequence of events when viewed quantitatively.

In 1510 there occurred in Europe (then, of course, the chiefly known portion of the world) a wave of influenza which many students regard as the first real outbreak which ever occurred. In the chronicles of earlier epidemics there is evidence of influenza outbreaks, but this is much confused because of the frequent recurrence of pestilences during the Middle Ages. Creighton quotes Short's statement that in 1510 "none died except some children"; Short failed to state his authority,

[1] The word "outbreak" is used synonymously with "epidemic" in this paper.
[2] For example, the outbreaks of 1767, 1781, 1802–3, and 1874–75.

but Creighton surmised it was probably the continental chroniclers. Therefore, as the mortality was doubtless low, we may suspect that the 1510 wave was only an incident in the decline of some great previous pandemic, perhaps that of 1173 (regarded by Hirsch as the first authentic epidemic) which was recorded in Britain, Italy, and Germany and doubtless appeared elsewhere on the Continent. The reason for suspecting this sequence would be its apparent consistency with the course of events in the subsequent history of the disease.

For practical purposes a pandemic epoch may be defined as beginning with that year in which those first authentic outbreaks occurred anywhere in the world which developed into waves of pandemic proportions, after a series of years in which apparently no epidemic occurred anywhere; the end of the epoch may be dated as either the last year of high (or severe) influenza mortality in Great Britain or of world-wide, or nearly world-wide, epidemic diffusion of the disease before a general decline of the mortality begins.

After the opening of the sixteenth century the first giant wave of influenza arose in 1557, rapidly spreading over the then known world. This outbreak persisted in Britain during 1558, taking a heavy toll of lives, and was followed by revivals up to the year 1580 both in Britain and on the Continent. This may be regarded as the first definitely known pandemic epoch. It covered a period of 24 years.

The second epoch began in 1729 and persisted to 1743, a period of 15 years, in the midst of which there occurred in Britain, in 1733 and in 1743, two very severe epidemics, which were believed by Creighton to be the most severe in the history of Britain previous to the time of his writing (1894). During this epoch, in various years, there occurred similar severe outbreaks in other countries, including North America. One writer stated that 80,000 persons were sick in Rome in 1743, and that 500 were buried in one day.

The third epoch began about 1824 and persisted to 1851, a term of 28 years, during which there occurred in different years and in various countries of the entire world very severe epidemics resulting in heavy mortality from the disease.

The fourth epoch may be regarded as that of 1889 to 1900, a term of 12 years, the details of which are recent knowledge.

The fifth epoch began in 1915 [1] and *is still in progress.* The first pandemic invasion was the great wave of 1918, which has been followed by several revivals in various countries, some of serious proportions as to mortality. This, doubtless, was the most fatal of all influenza pan-

[1] Influenza appeared widely in epidemic form in the United States in December, 1915, and January, 1916, being reported definitely from 22 states. See U. S. Public Health Reports for January 7, 1916.

demics and was perhaps the greatest single catastrophe of modern times, not excepting wars.

Therefore the duration of the first four epochs being roughly 24, 15, 28, and 12 years respectively, or an average of 19.8 years, it is clear that the epoch beginning in 1915 will very probably contain one or more severe epidemics, or pandemic waves, in addition to that of 1918. These outbreaks should be anticipated during the 10 to 20 years following 1915.

A curious fact in connection with each of these great epochs is that upon the occurrence of the first severe epidemic phase, both the laity and the medical world immediately questioned the identity of the infection with that of other previous outbreaks. In former times these epidemics were frequently referred to in Britain as epidemics of a "new disease," "strange disease," etc., and, indeed, even during the present pandemic period able observers have questioned whether the disease is the same as that of other years, or in one recent epidemic as compared with another. Regardless of clinical and pathological differences observed in the study of individual patients, a quantitative review of the mass of mortality data may be of value. Each of the five pandemic epochs just described was immediately followed by a series of years during which the mortality from the disease remained generally high. This mortality for each year is definitely known in England and Wales since 1838. In the outbreak of 1847 the influenza death rate per million was 285, in 1848 it was 460; and it remained above 75 for the next five years. In 1851 a recrudescence of the disease raised the rate to 193; the next year it was 55, after which it *declined continuously*, with minor variations, for many years until the decade 1880–89, when it was below 10. During the second half of that decade it fell to only *2 per million!* Then another pandemic invasion caused rates in 1890 and 1891 of 157 and 572 respectively. If depicted graphically (Chart II), this decline during the 40 years from 1850 to 1889 (or beginning approximately with the end of a pandemic epoch) makes a hyperbolic curve if smoothed. This seems to represent a natural phenomenon uninfluenced by human factors and obeying a law susceptible of mathematical expression.

During the interpandemic period from 1889 to 1918 the annual influenza mortality of England as a whole remained on a very much higher level than during the 40 years preceding 1889, and although characterized by extreme fluctuations it nevertheless presents a picture of general decline from 1890 to 1918. Perhaps the most striking and the most interesting feature of the estimated mortality of the epochs previous to 1838 is that each of them contains various minor epidemics

with moderately high mortality, in addition to very severe pandemic waves; and each such epoch is immediately followed, for a series of years, by a concentration of moderately high influenza mortality. After the severe outbreak of 1743 in Britain (which marked the end of a pandemic epoch) there occurred during the 19 succeeding years many minor epidemics of influenza and colds, with a well-marked effect on the general mortality, which corresponds to a similar condition observable in the official statistics after the epochs of 1824–51 and 1889–1900. Minor or secondary epidemics after the pandemic wave of 1918–19 were noted during 1919–22 in several countries.

After these periods of high mortality, which followed the great epochs, there still occurred in isolated years in Britain mild epidemics, occasionally very widespread, some of which are doubtful as to identity and some unmistakably real influenza. These, however, left but slight impress upon the bills of mortality and the later official statistics. Thus, as a whole, these periods fit into the picture as do the declining annual death rates in the Registrar-General's statistics during the 40 years before 1890. During the years from 1762 to 1809, when the very mild and unimportant outbreaks were disappearing, it will be noted that these *become less frequent* as the distance becomes greater from the pandemic epoch; so that in 1807–8 slight epidemics of "catarrhal" diseases prevailed in Britain, the character of which was doubtful and the identity with influenza disputed. *These mild outbreaks were followed by a term of 17 successive years during which there occurred no epidemic of any disease resembling influenza.* Then during the 6 years from 1826 to 1832, mild epidemics of a "catarrhal" type of sickness again appeared, and in 1833 there was a very serious influenza epidemic. These slight "catarrhal" outbreaks may be regarded as preliminary to the oncoming storm and as being related directly to the wide diffusion of mild influenza outbreaks in the United States in 1824–25, and in all the Western Hemisphere, Russia, and Siberia in 1826–27, which preceded the universal appearance of the disease in the giant waves of another pandemic epoch. A similar phenomenon preceded 1918, in the constantly increasing pneumonia mortality during several preceding years. One may well consider whether such apparently unimportant but constantly increasing respiratory mortality is not essentially the beginning phase of the pandemic and susceptible of statistical definition.

When the similar events of the great epoch of 1729–43, and its slow disappearance over many years to 1824, are charted, the picture presents features unmistakably like those in the origin of the next pandemic epoch, beginning in 1824, and its subsequent decline down to

1889. It would require but little imagination to construct curves which would represent roughly the long interpandemic declines.

From the foregoing it may be clear why the mild outbreak of 1510 (mentioned early in this paper) may be viewed as an isolated incident in a long curve of decline from a previous great wave. Analagous declines of the mortality from other epidemic infections over long periods of years will be recalled, *e. g.*, plague and leprosy.

One of the many other interesting observations which might be made as to the great influenza epochs is the *apparent* shortening of the time intervals between them during the last four centuries. This relative shortening will be found to exist, regardless of the exact years used to measure the pandemic epochs. These intervals, *or long-time cycles*, may be defined as follows, from the beginning of one epoch to that of the next:

From the epoch of 1557–80	to that of 1729–43	172 years
" " " " 1729–43	" " " 1824–51	95 "
" " " " 1824–51	" " " 1889–1900	65 "
" " " " 1889–1900	" " " 1915–(?)	26 "

We may supplement this interpretation of the statistical evidence of influenza with these judgments:

1. Pandemic influenza is identical with the influenza of the interpandemic periods: it is an intensely virulent and highly communicable form of the same disease. The disease may differ in clinical or pathological manifestations in various outbreaks.

2. Influenza exists constantly in mild or severe form in one or more countries of the world, or subdivisions of them; and factors, at present unknown, occasionally operate to increase the virulence and communicability of the disease, thus causing either epidemics or pandemics. (This condition becomes apparent when the influenza mortality of Britain is studied in comparison with the historic epidemics.)

3. There is increasing severity and frequency of local outbreaks in one or more countries for several years preceding the first invasion of a pandemic epoch; these preliminary outbreaks seem to begin after a series of years entirely free from epidemics, and may be susceptible of sanitary recognition, if not statistical demonstration.

4. The appearance of influenza in pandemic form is evidence of the presence of a pandemic epoch, during which period of years one or more other severe epidemics or pandemics will occur at intervals of a few months to a few years.

5. The rise and fall of the pandemic epochs, in long-time cycles, indicates that the fundamental causative factors probably have been uninfluenced by preventive measures.

THE NET VOLUME OF SAVING IN THE UNITED STATES

MEASURED BY THE INCREASE IN THE TOTAL WEALTH OF THE COUNTRY

By Willford I. King, *National Bureau of Economic Research*

Part II

F. AN INDEX OF THE VOLUME OF NET SAVINGS IN THE CONTINENTAL UNITED STATES

The purpose of this section is not to measure the absolute volume of savings in the country, but merely to obtain an index representing the relative changes from year to year in the total volume of the savings of the American people. This index is made up of three components, namely, business savings, personal savings, and governmental savings. Although these three categories are in many respects separate, they are frequently inversely correlated, for the credits in one class often represent debits in another. Thus personal savings and business savings may be invested in government bonds issued to obtain money for current expenses. In this case, although individuals and business firms are increasing their wealth, the government is evidently increasing its debt by an equal amount without accumulating any offsetting assets, and no net social saving has resulted. An index of total savings ought, therefore, to be made up of items representing respectively the three categories mentioned.

The estimate here presented of the savings made by the various business enterprises of the United States is based primarily upon the study recorded in Volume II, Section 24e, of *Income in the United States;*[1] but the totals are different because the estimates of agricultural savings have been revised. Of course, also, the amounts saved each year from the personal incomes of the inhabitants are very large. Since these latter amounts are not recorded directly they can only be estimated from the volume of investments recorded in certain fields.

Business enterprises which have adequate accounting systems, in calculating their surpluses for the year, always make an allowance for depreciation. If we are to compute the aggregate net saving made from personal income, it is similarly necessary to deduct from gross savings amounts which have been lost, wasted, or consumed during the period. As a matter of fact, the data available are not sufficiently complete to enable us to measure the volume either of gross or of net

[1] By the National Bureau of Economic Research.

personal savings. It is believed, however, that by combining the annual increases in the respective volumes of certain types of assets an index is obtained which shows with a moderate degree of accuracy the relative amounts of net savings from personal income occurring in different years. The items combined to form this index are:

1. The increase in savings banks deposits in the United States.
2. The value of new buildings constructed in the United States, less the loss to existing buildings arising through fire and depreciation.
3. Total increases in the net assets of all insurance companies.
4. The face value of new issues of bonds listed on the New York Stock Exchange.
5. The value of the net increase in the number of passenger automobiles owned in the United States.
6. The net increase in holdings by American individuals or firms of foreign securities or other property.

All of these items seem to represent not only gains in individual wealth but, in most instances, growth in social wealth as well. As savings banks gain in deposits they normally add to their investments in productive enterprises. If new building is more than sufficient to offset the deterioration of existing buildings, there is evidently a net increase in the national wealth. Larger assets of insurance companies mean, as a rule, larger loans to entrepreneurs for the purpose of expanding industrial equipment. The proceeds of new [1] bonds listed on the New York Stock Exchange are utilized for like purposes.

While Orientals put their personal savings largely into gold or jewels, Americans build better houses, lay out more beautiful grounds, buy more elaborate furniture, and purchase more expensive automobiles. We have no record of the value of furniture amassed, but we have at least an inkling of the volume of that saving which takes the form of new pleasure cars. An estimate of this quantity has therefore been entered in the record.

In addition to noting variations in the value of private wealth located in the United States it is necessary to consider changes that have occurred in the debits and credits of people abroad in their business dealings with our inhabitants. Before 1914 the people of the United States usually increased their debt to Europe each year. Since that date the reverse has tended to be true. The present worth of claims against Europe is of course a matter of much uncertainty. In constructing this index the simple course has been followed of counting them at their face value.

The inclusion of all of the above items in an estimate of personal sav-

[1] Bonds not issued for refunding purposes.

ing in the United States seems logical. Although it is evident that they are not entirely independent but overlap in many instances, there seems, nevertheless, no reason for believing that their sum may not be used to advantage as an index for measuring the changes from year to year in the volume of net personal saving. At first thought it appears that the accuracy of this index might be improved by increasing the number of items upon which it is based. Careful consideration, however, has shown that some ostensibly satisfactory criteria in reality either are not measurable or fail to meet our needs. Examples of items which have been rejected on such grounds are (1) new stocks listed on the New York Exchange, excluded because stocks are sold only rarely at par for cash, and in many instances have no par value, and (2) gains in demand deposits in banks, excluded because they generally represent business borrowing rather than personal savings.

By adding together the six included items a sum is arrived at which is believed to represent moderately well the changes in the volume of personal saving from year to year. If to this sum we add the increases in the business savings of the country and subtract from the sum the increases in the national debt, the remainder may be considered a reasonably good index of the total volume of net savings in the United States. It is of course evident that it is only an index and must be expected to vary widely from the absolute total of the national savings. Unfortunately, our estimates for business savings do not cover the years preceding 1909, and our records of foreign debits and credits for the years before 1914 are extremely crude; hence the volume of saving for earlier years must be estimated on the basis of personal savings and national debt only. Building estimates form one of the important items in personal savings. These, for the earlier years, have been founded upon very tenuous assumptions. It follows, then, that while the estimates for 1909 and the later years are believed to be fairly satisfactory, the figures presented for the years preceding 1902 must be considered as having only slight value; hence it does not appear desirable to carry the computations for that period any further than is done in Table II. The figures in column D of that table serve to show something of the movements of savings during the two decades preceding the year 1909.

The fact should be noted that column K represents four-fifths of the inverse [1] of the changes in the debt of the federal government. The statement is made in the form of credits in order to make it comparable with the items in the other columns, all of which are stated in positive form. The reason for recording a part rather than the whole of this

[1] All signs changed from plus to minus or vice versa.

TABLE II

CRITERIA REPRESENTATIVE OF NET SAVING IN THE UNITED STATES AS MEASURED IN DOLLARS OF THE GIVEN YEAR

A	B	C	D	E	F	G	H	I	J	K
Year	Index of total net saving (C+D+K)	Estimated total saving by business enterprises [2] (millions)	Savings from personal income (millions)							Four-fifths of the increase in net credits of the federal government [3] (millions)
			Total of recorded items E+F+G+H+I+J	Increase in net value of			Increased supply of		Increase in private American credits abroad [8]	
				Assets of all insurance companies [3]	Savings Bank deposits [4]	Bonds listed on the N. Y. Stock Exchange [5]	Buildings in the United States [6]	Passenger automobiles in use [7]		
1890	....	....	....	....	$99	$4,198	$798	$0	$−48	$+50
1891	....	....	$1,148	$72	94	191	841	0	−50	+17
1892	....	....	1,279	81	81	175	994	[1]	−52	+ 5
1893	....	....	921	71	18	139	747	[1]	−54	−37
1894	....	....	643	84	13	85	517	[1]	−56	−12
1895	....	....	781	93	80	104	562	1	−58	−26
1896	....	....	667	87	64	47	527	1	−60	−28
1897	....	....	661	105	79	88	447	2	−60	−32
1898	....	....	752	116	145	56	490	4	−60	−84
1899	....	....	1,056	143	192	156	618	6	−60	−11
1900	....	....	901	152	183	148	471	7	−60	+40
1901	....	....	1,184	168	150	220	696	9	−60	+54
1902	....	....	1,060	190	169	195	555	10	−60	+62
1903	....	....	1,106	183	155	192	627	9	−60	− 4
1904	....	....	1,457	243	163	356	731	24	−60	−36

1905	....	....	2,034	221	211	463	1,175	25	−60	+ 2
1906	....	....	1,777	228	214	300	1,058	36	−60	+76
1907	....	....	1,489	140	89	247	1,027	46	−60	−10
1908	....	....	1,821	347	12	658	788	76	−60	−73
1909	$4,707	$2,346	2,401	279	119	642	1,271	149	−60	−40
1910	5,071	2,778	2,294	245	209	538	1,138	224	−60	0
1911	3,973	1,938	2,021	294	171	391	1,029	196	−60	+14
1912	4,962	2,697	2,271	272	209	415	1,120	315	−60	− 6
1913	4,592	2,401	2,190	271	210	422	955	392	−60	+1
1914	3,937	1,414	2,552	283	136	352	785	431	+565	−29
1915	7,099	3,341	3,726	280	131	447	817	341	+1,710	+32
1916	11,810	6,966	4,985	387	377	236	1,178	690	+2,117	−141
1917	8,694	6,586	3,173	438	174	714	785	1,173	−111	−1,064
1918	−1,736	4,907	2,005	581	234	100	576	551	−37	−8,648
1919	....	....	3,935	291	506	226	1,399	1,087	+426	−257

[1] Less than one-half million.

[2] Estimated for non-agricultural industries from the amount of business savings in those industries for which records are available by using the ratio of the share of entrepreneurs and other property owners in all industries in this group to the same share in those industries recording business savings. The quantities thus estimated, when added to the savings in agriculture, give the figures in column C.

[3] Compiled from *Insurance Yearbooks* published by the Spectator Co.

[4] From the *Annual Report* of the U. S. Comptroller of the Currency, 1920, Vol. I, pp. 236–142.

[5] Compiled from the *Financial Review*. Foreign bonds and U. S. Government bonds except Panama Canal issues excluded. Foreign bonds appear in column J. The U. S. Government bonds excluded represented current expenses and not savings.

[6] A rough estimate for years previous to 1902; figures for later years based on records of building permits; depreciation and fire losses deducted.

[7] Computed on basis of figures in *Facts and Figures of the Automobile Industry*, by the National Automobile Chamber of Commerce, and in *The Automobile Industry and Its Future*, by Leonard P. Ayres.

[8] Estimated from studies in *Harvard Review of Economic Statistics*, 1919 to 1921.

[9] For description of this item and the reasons for using only four-fifths of the total, see the text.

quantity is that the use of the whole would presumably tend to give this item too much weight in the composite index. The sum of personal savings shown in column D is probably about four fifths of the actual total of personal savings in the United States, and column K has therefore been reduced in somewhat the same proportion to prevent relative overweighting. It is possible that the item of business savings should be similarly reduced, but the data in this connection are believed to be better criteria of national saving than are those entered in columns D and K, and hence no reduction has been made.

Table II makes it clear that while the volume of saving by business concerns varies directly with the waves of business activity, the extent of private saving is much less closely correlated with the economic cycle.

G. THE TOTAL ANNUAL SAVINGS OF THE PEOPLE OF THE UNITED STATES

At first thought it might seem that if the value of the wealth at the beginning of the year were subtracted from the value of the wealth at the end of the year, the net result would be the annual savings measured in terms of money. More careful consideration indicates, however, that because of the manner in which wealth is currently valued, a figure arrived at in this manner has little significance.

As previously pointed out, the current valuation of durable goods depends as much upon the views generally prevailing at the moment concerning the future course of prices, interest rates, etc., as it does upon any definitely known facts concerning present income or physical wealth. The value of existing wealth, therefore, varies greatly with every change of community opinion, with every wave of pessimism or optimism. It seems absurd, however, to say that because people have suddenly become more hopeful, the nation has thereby accumulated billions of savings. It appears, therefore, that estimates of the national wealth can be safely utilized only as starting points upon which to base estimates of national savings, and that a rather roundabout calculation is necessary to arrive at anything approaching a logical estimate of the savings of the country.

The calculation here presented is founded upon the following assumptions:

1. Saving should be measured as a differential between valuations derived by a constant mode of reasoning from trends of income and should not be calculated upon the basis of merely momentary conditions and opinions.

2. The present value of wealth depends upon anticipations concerning income to be derived from present wealth in the future and not

at all upon estimates of income to be obtained in the future as payment for personal services.

3. The average ratio between wealth and income from property and entrepreneurial efforts over the period 1909 to 1915 inclusive probably furnishes a moderately good criterion of the ratio which the beliefs and feelings of the community normally fix between these quantities. Greater accuracy can be secured when records for more years become available. The ratios for 1916 to 1918 as shown in Table III, column D, have not been used in computing the average, since it appears that the unusual price inflation made them abnormally low.

4. Total national saving probably bears an approximately constant ratio to an index of business, private, and public savings combined.

The principles just enunciated are exemplified in Tables III and IV. All values have been reduced to the price level of 1913 in order to eliminate fictitious increases in wealth due to changes in the value of money.

The purpose of Table III is to ascertain the total national saving for the nine-year period, 1909–18, as measured in terms of dollars of 1913. Since the normal ratio of the current value of all property in the United States to the annual income of entrepreneurs and other property owners seems to be around thirteen to one, it follows that if currency changes affect about equally the values of direct and indirect (consumption and production) goods, the wealth of the country, when converted into terms of constant purchasing power, ought to be approximately thirteen times the income of the above-mentioned classes, when such income is expressed likewise in terms of constant purchasing power. It is indeed true that for any given year a ratio based upon physical quantities of wealth and physical quantities of income probably would vary greatly owing to differences in climatic conditions or to disturbances in industrial conditions. However, it seems probable that a moving average of such ratios calculated on a three or four year period would fluctuate but slightly. It is such a ratio that we are attempting to approximate by averaging the ratios in column D for the years 1909 to 1915.

Evidently the application of this average ratio to the purchasing power income for different years will scarcely give results for individual years which possess much merit, but the chances are that the error in any given year will be too small to affect materially the total difference between the respective wealth estimates for the end of 1909 and the end of 1918. The figure $46,115,000,000, calculated in the manner described,[1] may then prove to be fairly close to the actual increase in

[1] See Table III, column F.

TABLE III

A PRELIMINARY ESTIMATE OF THE WEALTH OF THE UNITED STATES MEASURED IN DOLLARS HAVING THE SAME PURCHASING POWER AS IN 1913

A	B	C	D	E	F
Year	Total income of entrepreneurs and other property owners [1] (millions of current dollars)	Wealth at close of year in current money [2] (millions)	Ratio of C to B	Purchasing power of total income of entrepreneurs and other property owners at prices of 1913 [1] (millions)	Preliminary estimate in terms of dollars of 1913 of the wealth at the close of the year. (E times the average ratio of C to B) (millions)
1909	$14,004	$172,125	12.291	$14,635	$188,674
1910	14,901	183,944	12.344	15,234	
1911	14,099	185,922	13.187	14,319	.....
1912	15,390	195,200	12.684	15,457	.....
1913	15,634	201,271	12.874	15,634	.....
1914	14,849	203,048	13.674	14,699	.....
1915	16,565	218,571	13.195	16,359	.....
1916	22,303	246,139	11.036	20,253	.....
1917	26,869	272,147	10.129	20,830	.....
1918	27,355	294,145	10.753	18,212	234,789
1919	....	....	....	....	.....
	Average for 1909 to 1915 inclusive = 12.892			Increase in wealth in nine years = $46,115	

[1] Revision according to information at hand on June 25, 1922, of estimates presented by the National Bureau of Economic Research in *Income in the United States*.

[2] See Table I, last column, this JOURNAL, September, 1922, p. 322.

wealth during the nine year period, if such wealth is, in each year, measured in terms of 1913 dollars. Since this estimate is the figure upon which all later results depend, it is manifestly worth while to test its accuracy with as much care as possible. It represents an increase in wealth of about 24.4 per cent, which increase is equivalent to about one seventh of the total income during the nine years. This seems to be a surprisingly large fraction. Is it not probable that it is so large merely because of some error in the basic figures?

Manifestly, if the calculated increase in wealth, as shown in column C of Table III, is greater than the facts warrant, it will make the difference between 1909 and 1918 too great. As a matter of fact, nothing in the mode of deriving the indexes used in computing Table I causes us to suspect that their slope is materially at fault.

But even if the slope of the items in column C is correct, if the absolute size of the items throughout is exaggerated the distance between the extremes will be exaggerated in the same degree, and an error will arise from this source. As a matter of fact, it appears more probable that the Census of Wealth, in compiling property values, has omitted many very valuable items, such as parks and other public property, from its totals than that it has duplicated other items of value. The chances are, then, that the volume of saving is not exaggerated because of this reason.

If the ratios shown in Table III, column D, for the last three years are included in the average, that figure will be reduced to approximately 12.2 instead of 12.9. Such a change would affect the results by only about 5 per cent. Evidently errors from this source cannot be considered of major importance.

Although the evidence at hand shows no reason for doubting that 46.1 billions of dollars of 1913 value represents approximately the extent of the saving during the nine years, it nevertheless seems worth while to test the accuracy of this figure by another set of computations. If wealth has increased at any such rate some evidence of the fact should be shown not only by the increase in the value of property rights represented by securities, but also by the increase in those tangible physical assets, the numbers or quantities of which have been recorded from time to time.

Table IV is made up of a collection of data which indicate the changes in the quantity of a limited number of items which are used as criteria because they are the best available. It is evident that they do not cover all of the important categories of wealth and that the units are not exactly comparable in nature at the two dates.

It is probable, however, that while the quality of some units—for

TABLE IV

FRACTION OF INCREASE DURING NINE OR TEN YEARS IN THE SIZE OF ITEMS REPRESENTING THE VOLUME OF PHYSICAL WEALTH IN THE CONTINENTAL UNITED STATES

Kinds of wealth	Beginning of period		End of period		Weight [12]	Product in millions	
	Year	Amount	Year	Amount		Beginning of period	End of period
Acreage of all crops (millions)[1]	1909	311.3	1918	365.2	70	21,791	25,564
Improved farm acreage (millions)[2]	1910	478.5	1920	507.	30	14,355	15,210
Woodland in farms (millions of acres)[2]	1910	190.9	1920	168.6	20	3,818	3,372
Farm implements and machinery [3]	1909	1,172.9	1918	1,800.7	3	3,519	5,402
Total farm land, machinery and agricultural products	....	....	....	....	....	43,483	49,548
Poultry on farms (millions)[2]	1910	295.9	1920	372.8	1	296	373
Horses, mules and cattle (thousands)[2]	1910	91,179.	1920	96,600.	90	8,206	8,694
Sheep, goats and swine (millions)[2]	1910	115.3	1920	101.	9	1,038	909
Total live stock	....	....	....	....	....	9,540	9,976
Gold coin and bullion in Treasury (millions of dollars)[4]	1909	1,642.	1918	3,076.	1	1,642	3,076
Freight cars (thousands)[4]	1909	2,173.	1918	2,516.	2,500	5,432	6,290
Locomotive tractive power (millions of pounds)[4]	1909	1,504.	1918	2,221.	1	1,504	2,221
Railways (thousands of miles)[4]	1909	244.1	1918	264.2	15,000	3,661	3,963
Passenger cars in service on railways [4]	1909	45,584.	1918	56,505.	8,000	365	452
Total, railways and equipment [5]	....	....	....	....	....	10,962	12,926
Telegraph wire (thousands of miles)[4 5]	1907	1,578.	1917	1,890.	100	158	189
Ocean cable (miles)[4 5]	1907	46,301.	1917	71,251.	2,000	93	143
American Telephone Company wire operated (thousands of miles)[5 6]	1909	9,831.	1918	22,610.	150	1,475	3,391
Electric railways (miles)[6]	1909	40,490.	1918	48,484.	80,000	3,239	3,879
Electric railway cars (number)[6]	1909	91,153.	1918	101,734.	20,000	1,823	2,035
Primary electric horsepower (thousands)[7]	1907	618.	1917	17,137.	260	1,721	4,456
Total public utilities	....	....	....	....	....	8,509	14,093
Horsepower used in factories of selected states (thousands)[8]	1909	1,207.	1919	1,768.	30,000	36,210	53,040
Number of motor cars registered (thousands)[9]	1909	300.	1918	6,147.	400	120	2,459
Tonnage of merchant marine (thousands) [4]	1909	7,389.	1918	9,925.	70	517	695
Navy tonnage in commission (thousands)[10]	1910	1,097.	1918	1,754.	800	878	1,403
Pupils in public schools (thousands)[4 5]	1909	17,506.	1918	20,854.	200	3,501	4,171
Mileage of improved roads (thousands)	1909	190. [13]	1919	296.	6,000 [14]	1,140	1,776
Standing timber (billions of board feet)	1911	3,243. [15]	1920	2,214. [16]	.0015	4,865	3,321
Property rights in foreign countries [11]	1909	−1,740.	1918	2,900.	1	−1,740	2,900
Total						$119,627	$159,384
Index—base, beginning of period						1.000	1.332

FOOTNOTES TO TABLE IV
ON PAGE 464

[1] *Yearbook of Agriculture*, 1920, p. 808.
[2] U. S. Census; Preliminary Reports for 1920.
[3] Value in millions of dollars at prices of 1914; computed from Census figures.
[4] *Statistical Abstract of the United States*, for 1920, pp. 325, 335, 367, 594, 811, and 818.
[5] Used as a measure of the total supply of equipment.
[6] *Statistical Abstract of the United States* for various years.
[7] United States Census of Central Electric Light and Power Stations, 1917, p. 49.
[8] United States Census of Manufactures, 1909, Vol. VIII, p. 511 and the 1919 Census Reports for sample States.
[9] *Automotive Industries;* January 15, 1920, p. 111.
[10] *Annual Reports* of the Secretary of the Navy.
[11] Items in Table I divided by indexes of wholesale prices.
[12] Weights have been chosen with a view to making the totals for the given groups of items roughly equal to the total value of this variety of wealth in 1918 as estimated in Table I. The weights assigned to the respective items indicate the relative importance attached to each as an indicator.
[13] United States Office of Public Roads, Bulletin 41, p. 41.
[14] *Statistical Abstract of the United States* for 1918, p. 297.
[15] Estimated from United States Forest Service Bulletin, *Our Timber Supply*, 1914.
[16] Greeley, William B., *Timber Depletion and the Answer*, United States Department of Agriculture, Circular 112.

example, freight cars—has improved, others like woodland may have grown poorer. The errors are, then, presumably to a large degree compensating.

Unfortunately, data are, in some instances, not available for 1909 and 1918, the two specific years desired. It has therefore been necessary in a number of cases to take in the decade 1909 to 1919. The inclusion of the extra year will of course make the computed percentage of increase for the period slightly too large.

The weighted average obtained by this method indicates a wealth increase for the period of about 33 per cent. If this be reduced by about one-fifteenth to allow for the error just mentioned, the estimated increase is reduced to 31 per cent, which is even larger than that obtained by the operations recorded in Table III. True, the weights used are necessarily more or less arbitrary, and experiment shows that a shift to other weights which seem almost as logical may change this result by as much as 2 or 3 per cent. Even after allowing for this degree of divergence, the estimate that the people of the country actually saved $46,115,000,000 in the nine years ending December 31, 1918, still seems more in danger of being too low than too high.

It is undoubtedly true that much saving can be accomplished without abstinence, for in the case of very wealthy persons it may be easier to save a large fraction of the income than to spend it all. However, such saving plays but a minor rôle in the national total. Certainly, almost anyone could spend $50,000 a year without difficulty. The total income of persons receiving more than that amount annually was, however, only $2,730,000,000 in 1918.[1] If we deduct from the sum just mentioned $50,000 for each of the persons in the higher income classes, the remainder will be only $657,000,000, or less than a seventh of the normal annual savings of the people of the nation. Although, then, much of the actual saving may cause relatively little sacrifice, it is evidently not true that any considerable proportion of the huge accumulation in the nation comes about simply because it is easier to save the income than to spend it.

Having estimated the total savings of the nation for the decade, the problem still remains of apportioning this sum among the various years of the period. No accurate method of accomplishing this end is available, but it seems probable that if the amount is assumed to be a constant multiple of the index of saving derived in Table II, the error in the resulting products will not be unduly large. This process has been carried out in Table V.

Table V brings out some interesting facts. It appears that the

[1] See *Income in the United States*, Vol. I, Table 27.

TABLE V

AN ESTIMATE OF THE NET SAVINGS OF THE UNITED STATES AND ITS PEOPLE

A	B	C	D	E	F	G	H
Year	Total national income measured in 1913 dollars (millions)	Total national income measured in dollars of given year [1] (millions)	Index of saving in the United States [2]	Ratio of income saved to index of saving [4]	Total national savings measured in dollars of current year (millions) (D×E)	Fraction of income saved $\frac{F}{C}$	Total income saved measured in dollars of 1913 (millions) (B×G)
1909	$30,608	$29,259	$4,707	1.0624	$5,001	.1709	$5,231
1910	32,176	31,471	5,071	1.0624	5,387	.1712	5,509
1911	31,397	30,904	3,973	1.0624	4,221	.1366	4,289
1912	33,985	33,807	4,962	1.0624	5,272	.1559	5,298
1913	35,436	35,436	4,592	1.0624	4,879	.1377	4,880
1914	33,067	33,401	3,937	1.0624	4,183	.1252	4,140
1915	35,150	35,920	7,099	1.0624	7,542	.2100	7,381
1916	41,669	45,861	11,810	1.0624	12,547	.2736	11,401 [7]
1917	42,377	54,664	8,694	1.0624	9,237	.1690	7,162 [7]
1918	38,830	59,930	−1,736	1.0624	−1,844	−.0308	−1,196 [7]
Total 1910–18	$324,095	$361,394	$48,402	Period	1909–18	.1561 [5]	$54,095 [6]

$\frac{\text{Savings for 1910–18 at 1913 prices}}{\text{Income for 1910–18 at 1913 prices}} = \frac{\$46,115}{\$324,095} = 0.14229$ = Fraction of income saved.

0.14229×$361,394 = $51,423 = Total income saved, in millions of current dollars.

$\frac{51,423}{48,402} = 1.0624$ = Ratio of income saved to saving index.

[1] Revision, according to information at hand on June 25, 1922, of estimates presented in *Income in the United States*.
[2] See Table II, column B.
[3] See Table III, column F.
[4] See computation at foot of this table.
[5] $\frac{\text{Sum of F}}{\text{Sum of C}}$.
[6] Total of above items.
[7] See the text for a discussion of possible adjustments in these items.

normal fraction of the national income saved is about one-seventh;[1] but that in a year of exceptional business prosperity like 1916 this fraction increased to one-fourth, while the war expenditures of 1918 actually used up not only all of the current income but also a very small fraction of the previously existing wealth of the country. If we measure values in terms of dollars of 1913 purchasing power, the saving of the country, during the decade under discussion, amounted to about $55 annually for each inhabitant of the United States.

Most readers will, presumably, not be surprised to learn that there was no net saving in the war year of 1918, but they may be inclined to doubt that the people of the nation actually added to their wealth in 1916 more than 12½ billions of dollars. Is such a figure reasonable? This question can best be answered by studying the statistics of total income measured in dollars of uniform purchasing power.

The income of all the employees in the Continental United States had been averaging about 19 billions a year for several years preceding 1916. In that year it increased suddenly to 21½ billions. If the savings of this class in ordinary years amounted to one billion dollars, it seems that in 1916 they may easily have laid aside 20 per cent of this extra pay, thus arriving at a total of 1½ billions.

Apparently, in the earlier years of the decade, entrepreneurs and other property owners had been in the habit of spending about 10½ billions and saving about 4½ billions a year. If in 1916 they spent 10½ billions as usual, there still remained nearly 10 billions for savings. In this case the 11.4 billions of savings would be approximately accounted for.

Unfortunately, we have no accurate evidence concerning the respective total savings of entrepreneurs and employees; but if the line of reasoning given above is correct, some interesting deductions follow. If we speak in terms of 1913 prices throughout, the 27,000,000 employees probably saved during the decade 1909–18 from their total income of 189 billions of dollars some 10½ billions of dollars or an annual average of $40 each, while each spent some $650 a year to support himself and his dependents. During the same period, the 6,400,000 farmers of the country and an unknown number of urban owners of farm land and farm mortgages received together some 54 billions of dollars. The average farmer's income must, then, have been noticeably below $850 per annum.[2] Out of the net value product of agriculture, the savings amounted to about 6,713 millions of dollars [2] or over $100 a farmer per annum.

[1] The fact should be kept in mind that this total includes additions to the stock of consumption goods as well as improvements in business equipment.

[2] Recent estimate by the National Bureau of Economic Research.

The average farmer therefore spent for himself and his dependents something materially less than $750 per annum, a sum which, however, in comparison with the average income of other entrepreneurs is not quite so meager as it appears, since a considerable share of the farmer's income consists of farm products consumed at home and these farm products are valued at farm prices which are on the average not more than half as high as the urban resident must pay for the same articles.

Aside from the farmers there are in the country about 3,300,000 entrepreneurs, and these receive on the average a total annual income of a trifle over 11 billions of dollars, or nearly $3,400 each per annum.[1] Out of this total these entrepreneurs probably saved, either directly or through their business enterprises, some 37 billions or slightly over $1,100 per man annually, leaving something under $2,300 each which they spent yearly for the consumption goods used by themselves and their families.

TABLE VI

THE WEALTH OF THE UNITED STATES, MEASURED IN DOLLARS HAVING THE SAME PURCHASING POWER AS IN 1913

Computed according to mathematical rules with unchanging principles of valuation applied throughout. Base, January 1, 1912

A	B	C	D
Beginning of year	Wealth at beginning of 1912 measured in dollars of 1913 value [1] (millions)	Net saving during year measured in dollars of 1913 value [2] (millions)	Wealth at beginning of year measured in dollars of 1913 value [3] (millions)
1909	.....	$5,231	$172,380
1910	.....	5,509	177,611
1911	.....	4,289	183,120
1912	$187,409	5,298	187,409
1913	.....	4,880	192,707
1914	.....	4,140	197,587
1915	.....	7,381	201,727
1916	.....	11,401	209,108
1917	.....	7,162	220,509
1918	.....	−1,196	227,671
1919	.....	.....	226,475

[1] When computed from the purchasing power of the total income, it appears that the value in current money of the wealth on January 1, 1912, should be multiplied by 1.008 in order to arrive at the value of the wealth on January 1, 1912, in terms of the prices of January 1, 1913.
$185,922,000,000×1.008 = $187,409,376,000.
For data see Table I and columns B and C of Table V.
[2] See Table V, column H.
[3] Computed by addition or subtraction from columns B and C.

If the estimates of saving shown in Table V are accurate, and if the Census gives the correct total value for the national wealth in 1912, then the items in Table VI cannot be materially in error.

Although it must be admitted that the figures in the last column of

[1] See Table III, column E.

Table VI do not correspond to customary ideas of value, it is believed, nevertheless, that they represent as closely as the accuracy of the data will permit an index of the increase during the decade in wealth, if wealth is thought of rather as a mass of goods than as an aggregate of values. When viewed from this standpoint it appears that wealth grew steadily throughout the decade, with the exception of the last year considered, and that the total at the end of the period was about one-third greater than that at the beginning.

TABLE VII

A COMPARISON OF TOTAL SAVING WITH SAVING BY BUSINESS CONCERNS

For the Continental United States

A	B	C	D
Year	Total savings of nation [1] (millions)	Business savings [2] (millions)	Per cent of total saved by business concerns $\frac{100C}{B}$
1909	$5,001	$2,346	46.91
1910	5,387	2,778	51.57
1911	4,221	1,938	45.91
1912	5,272	2,697	51.16
1913	4,879	2,401	49.21
1914	4,183	1,414	33.80
1915	7,542	3,341	44.30
1916	12,547	6,966	55.52
1917	9,237	6,586	71.30
1918	−1,844	4,907	[3]

[1] See Table V, column F.
[2] See Table II, column C.
[3] Business savings are more than total, which is negative.

The calculations recorded in Table VII indicate that the business enterprises of the country are normally responsible for about 40 per cent of the entire saving of the nation and, as might be expected, that this percentage increases in years of business prosperity and diminishes in times of depression.

LAW OF GROWTH IN FORECASTING DEMAND

By Raymond B. Prescott

A number of years ago in plotting the growth of the population in the United States as a whole, and also in several of the older states of the Union, I found that they all traced a similar curve. While reading Tarde's *Laws of Imitation* I noticed that he mentioned a similar curve for the growth of railroads and of tobacco consumption in France. This led me to plot the growth of several products in the United States. Again I found a similar curve. I noted also that all the data plotted were of products that were a direct function of population. These results seemed to indicate that there is a general law of growth that would fit the growth of population, and also some products, perhaps all, that are a direct function of the growth of population.

The curve seems to pass through four phases, indicating that population and industry perhaps pass through the same phases or periods in the course of their development. An investigation of the growth of several industries seemed further to confirm this idea. The four stages seemed to be[1]:

(1) Period of experimentation.

(2) Period of growth into the social fabric.

(3) Through the point where the growth increases but at a diminishing rate.

(4) Period of stability.

If this growth is fundamental and common for all industries whose products are a direct function of the growth of population, and if this growth could be expressed by a general type of equation, it is not difficult to conceive that when one part of the growth is known the others can be determined. To explain past performances and predict the future are not different operations. They are the same operations working in opposite directions, one from effect to cause and the other from cause to effect.

During the period of experimentation the permanent success of an industry is doubtful because man, being a creature of habit, dislikes to have his routine changed. He accepts very slowly anything that tends to alter his course from its accustomed path; consequently he opposes the object that is tending to vary his usual habits. The success of the industry depends upon how well the resisting habits of man are

[1] I am indebted to Mr. Harry Tipper, of *Automotive Industries*, for aid in defining the four stages.

overcome. Some individuals are more easily overcome than others, and soon become part of the force working to overcome the general resistance.

If the industry succeeds in overcoming man's natural resistance to having his habits disturbed, then progress of the industry enters into the period of growth into the social fabric, where it either becomes a part of the daily life of man or fails. Rate of growth varies with every industry, the rate depending upon the industry's adaptability to the economic and social life of man, which in turn depends upon the characteristics of the product and the incentive for its use created by man himself.

During the course of its progress an industry finally reaches a point where its growth continues to increase but at a decreasing rate, and so on into the period of stability where it becomes part of the economic and social life of man.

Merely to know that such growth is a possible fact is not enough. To make this knowledge really useful it is necessary to measure this growth. Before this can be done it is necessary to have a clear conception of what constitutes and influences this growth.

Every individual exerts some influence upon the growth of an industry. These influences vary in intensity and quality, from an extremely favorable attitude to one strongly unfavorable. The yearly summation of these favorable or unfavorable influences determines the progress of an industry. Some of these influences are more vital in shaping the course of the industry than others; some are so slight that they hardly make any impression at all. It is the combination of these many influences each year that finally determines whether or not the industry grows.

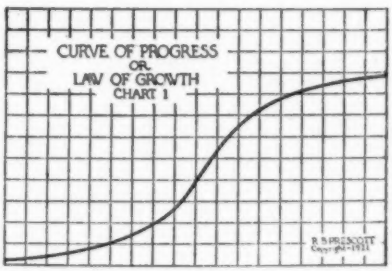
These influences may be grouped under two heads, "internal" and "external." The internal can be controlled to a large extent, and the external can be created by the proper action of the internal influences. If the internal influences are very strong, they naturally create more favorable external influences, sometimes called "good-will."

Strong internal influences are created by having a strong organization, good products, and sound policies. Of course there are many factors that enter into these broad divisions which permit of many combinations of favorable or unfavorable influences. One of the most common is the personality of the executives upon whose shoulders lies the responsibility for selecting the members of an organization, developing the product, and creating the policies under which the industry is operated.

In a large industry some companies make real progress while others

remain stationary. The successful company is one in which internal influences have been combined, resulting in a strong organization with a concomitant creation of favorable external influences. The degree to which the favorable external influences are created depends entirely upon the extent to which the favorable internal influences dominate the combination.

If an organization has a good product that is properly manufactured, advertised, and sold, a demand will naturally develop for the product and good-will will be created in the minds of users. In order that an industry may make steady progress under our competitive system, it must continue to have strong internal influences which are constantly creating strong external influences; otherwise it would retrogress.



The result of the summation of these many influences is measured in the yearly volume of business transacted, whether it is in physical units or in value. The geographical expression of this growth is shown by the curve traced in Chart 1. This curve seems to be fundamental for all industries that are either directly or indirectly functions of population; *i. e.*, whose growth depends directly or indirectly upon the ability of the people to consume their products. Every country may have a different rate of growth, and so may every industry, because no two industries have the same combination of influences. They will trace the same type of curve, however, even though the rate of growth is different.

It seemed as though it might be possible to obtain a general equation by which one could express and measure this phenomenon of growth as shown in Chart 1. After I had experimented with several equations that would fit the conditions for part of the curve, the Gompertz curve[1] was suggested to me. I experimented with this curve, first with population and then with a number of products, with quite satisfactory results, which indicated that it might be possible to forecast the probable growth of an industry and, by so doing, the probable demand for the product. Of course it must be realized that even if this equation will forecast probable or normal demand for a product, the business cycle must be taken into account before the actual demand can be

[1] I am indebted to Professor Harry C. Carver, of the University of Michigan, for suggesting the Gompertz equation as a curve having the desired trace, and to Professor Henry L. Moore, of Columbia University, for further assistance in its use.

forecast for any particular year. I believe that this equation can be used to forecast the normal growth of an industry over a period of time, though the yearly functions may be above or below this curve. It is the average of these variations that determines the growth.

The equation[1] is as follows:

$$y = ab^{c^x}. \tag{1}$$

$$\text{Log } Y = \text{Log } A + c^x \text{ Log } B. \tag{2}$$

Let now x become successively x_1, $x_1+1+ \quad . \quad . \quad . \quad +x_1+n-1$. Then taking the sums we get

$$\sum_{x}^{x+n-1} Y = nA(c^x + c^{x+1} + \quad . \quad . \quad . \quad + c^{x+n-1})B. \tag{3}$$

The series of c is in geometrical progression.

Such a series is:

$$S = a + a\beta + a\beta^2 + \quad . \quad . \quad . \quad + a\beta^{n-1}$$

$$\therefore \beta S = a\beta + a\beta^2 + \quad . \quad . \quad . \quad + a\beta^{n-1} + a\beta^n$$

$$\therefore S(1-\beta) = a(1-\beta^n) \text{ and}$$

$$S = a\frac{(1-\beta^n)}{(1-\beta)}.$$

Consequently the sum of the c series is

$$S = c^x\frac{(1-c^n)}{1-c} = c^x\frac{(c^n-1)}{(c-1)},$$

and we may write (3)

$$\sum_{x}^{x+n-1} Y = nA + c^x\frac{(c^n-1)}{(c-1)}B. \tag{4}$$

This equation (4) will enable us to find the constants in the formula.

There are three constants in the equation, and we therefore need to have three simultaneous equations. If we divide up our observations into three equal parts we have, using equation

$$\sum_{x}^{x+n-1} Y = nA + c^x\frac{(c^n-1)}{(c-1)}B \tag{4}$$

$$\sum_{x+n}^{x+2n-1} Y = nA + c^{x+n}\frac{(c^n-1)}{(c-1)}B \tag{5}$$

$$\sum_{x+2n}^{x+3n-1} Y = nA + c^{x+2n}\frac{(c^n-1)}{(c-1)}B.$$

[1] As far as I have been able to learn, this equation has never been fitted to economic data before.

Differencing both sides of these equations, we get

$$\sum_{x+2n}^{x+3n-1} Y - \sum_{x+n}^{x+2n-1} Y = c^{x+n}\frac{(c^n-1)^2}{(c-1)}B$$

$$\sum_{x+n}^{x+2n-1} Y - \sum_{x}^{x+n-1} Y = c^{x}\frac{(c^n-1)^2}{(c-1)}B.$$

These two equations may be written

$$(6)\qquad \begin{aligned} \Delta \sum_{x+n}^{x+2n-1} Y &= c^{x+n}\frac{(c^n-1)^2}{(c-1)}B \\ \Delta \sum_{x}^{x+n-1} Y &= c^{x}\frac{(c^n-1)^2}{(c-1)}B. \end{aligned}$$

Now dividing both sides of the first equations in (6) by the corresponding side of the second equations in (6) we get

$$\frac{\Delta \sum_{x+n}^{x+2n-1} Y}{\Delta \sum_{x}^{x+n-1} Y} = c^n.$$

Take the logarithms of both sides

$$(7)\qquad \log \Delta \sum_{x+n}^{x+2n-1} Y - \log \Delta \sum_{x}^{x+n-1} Y = n \log c.$$

Equation (7) enables us to compute at once the value of c. We may then find (b) from either of the equations in (6). Finally A may be obtained from (2).[1]

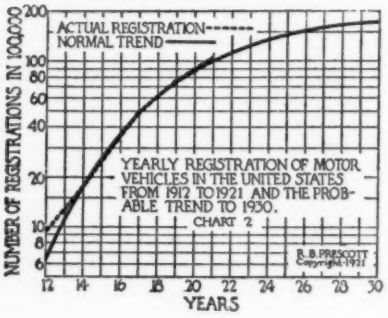
In studying the factors which influence the law of growth of populations or of an industry that is a direct function of population, motor vehicle registration data were fitted to the previously mentioned equation, as I was particularly interested in the automotive industry.

It was thought that if results obtained by fitting these data to the equation were reasonably satisfactory, it would create possibilities for reaching the solution of many problems that face the industry, the principal problem being to determine the number of motor vehicles that can be absorbed in the United States yearly.

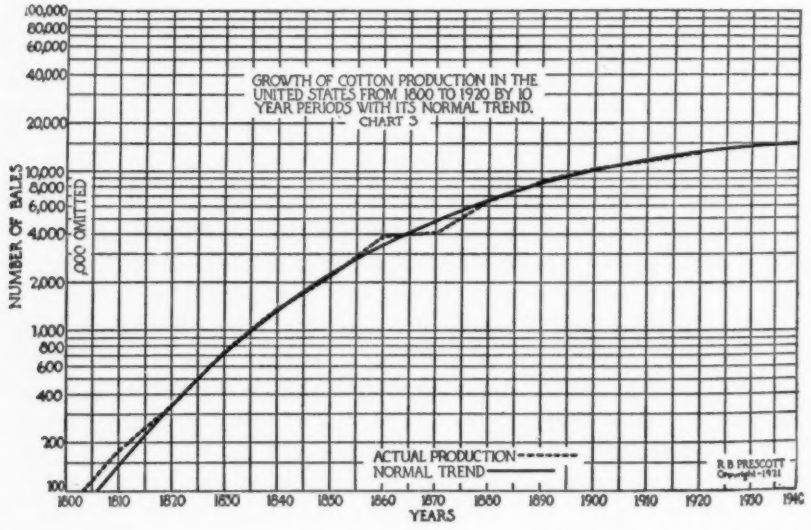
The automotive industry is of comparatively recent origin; hence data for only a few years are available, and the results of the research probably cannot be considered to have the same degree of accuracy in measuring the law of growth, or in forecasting over as long a period, as they would have if data for an older and more firmly established industry, which has entered the period of stability, had been used. The

[1] The equation was first published in *Automotive Industries*, Nov. 17, 1921.

results of the experiment may be seen in Chart 2 which shows the number of vehicles registered yearly, with the year 1921 as a forecast. The curve is extended to 1930 for the purpose of seeing how soon motor vehicle registrations will approximate the same rate of growth as population. It is not my intention that there should be considered seriously any forecast of more than a year or two, or perhaps three years beyond the data used, which in this case was up to 1920.


ACTUAL REGISTRATION
NORMAL TREND
YEARLY REGISTRATION OF MOTOR VEHICLES IN THE UNITED STATES FROM 1912 TO 1921 AND THE PROBABLE TREND TO 1930.
CHART 2
R.B. PRESCOTT
NUMBER OF REGISTRATIONS IN 100,000
YEARS

To show that with more extensive data it is possible to determine better the law of growth and to forecast more accurately for a longer period than with the data over a short period as in the case of the automotive industry, an older industry was used. For the experiment the total production of cotton was used for each tenth year beginning with 1820 and ending with 1890. The available data were then fitted to the above equation and the curve


GROWTH OF COTTON PRODUCTION IN THE UNITED STATES FROM 1800 TO 1920 BY 10 YEAR PERIODS WITH ITS NORMAL TREND.
CHART 3
ACTUAL PRODUCTION
NORMAL TREND
R.B. PRESCOTT
NUMBER OF BALES
000 OMITTED
YEARS

plotted as shown in Chart 3. The curve was then extended to 1920, a period of thirty years beyond which any data were assumed to have been available, with the pleasing results of only about a 2 per cent error in 1920. The forecasts for the three years were as follows:

	Actual	Graduated
1900.....	9,863,000	10,123,000
1910.....	11,600,000	11,608,000
1920.....	13,261,000	12,987,000

These results may be just a coincidence; but it hardly seems probable that three years, ten years apart, could be forecast as accurately and be merely a coincidence.

Many of the organizations in the industries that have kept any statistics have confined them almost entirely to their respective companies. That is all right as far as it goes, but it is only the beginning. Data showing only the yearly growth of an organization with respect to itself give li tle information as to what the real demand for the product is, or what percentage of the total business it is securing. An organization must know its progress with respect to the whole industry and what percentage of the total yearly business it has been getting over a period of years in order to obtain the proper perspective of its place in the industry. Such information will permit the executives of an organization to see the true position of the company and to determine its growth with respect to the whole industry.

To apply this equation to a single company would be just as dangerous as it would be for an insurance company to try to insure each individual without any knowledge of the law of probability. In an industry composed of a number of companies there is an opportunity to use the law of large numbers, which large number is the summation of their yearly volume of business. By use of these numbers an average can be forecast, and the probable error of variation and the percentage of the total business that a company has had over a period of time can be determined. On that basis the percentage of business that a company will secure for a year or two in advance can be forecast within probable limits of variation. This variation can no doubt be decreased if the demand for the product of any industry is adjusted to the business cycle.

By this method of forecasting the possible demand an industry can also determine the yearly probable demand for its product, and the various companies that constitute the industry can determine their respective positions. They should then confine their efforts to supplying that total, and should not over-expand, as has been the custom of the past. If this condition could be realized, it probably would eliminate some of the variation in the business cycle and make the probable error in forecasts much smaller.

It is hoped that large industries will realize that in forecasting the demand for a product for a whole industry there is a maximum that the

country can consume yearly, and that as the price increases, the possible demand becomes more and more limited. It has been found that by dividing the sales of cars into price classes, it is possible to forecast the probable percentage of the total sales in each group. For example, the probable sales for 1921 in each price class were forecasted in February of 1921; at the end of the year the actual sales were determined and the following comparison was made:

	Under $1,000	$1,000–$1,999	$2,000–$2,999	$3,000–$3,999	$4,000 and over
Calculated percentage	75.0	19	4.0	1.0	1
Actual percentage	73.5	20	4.2	1.5	1

Professor H. L. Moore, of Columbia University, published in the March issue of this Journal an equation which probably will give a much more refined method of forecasting the number of cars that can be sold in each price class, or as a matter of fact, at any price. I am at present experimenting with this method as applied to automotive data. It is too early to say what the results will be.

In a period of time the production of every company in a large industry bears a certain relationship to the total production, and the yearly rates of its business compared to the total business of the industry over a certain period of time should give that company a clear idea of its relative position and the progress it has made.

Let us assume that a company has acquired the total production of its specific industry and has determined its yearly rate of business over a period of time, the more years the better. It then uses this growth curve or some other method to forecast the probable demand for the next year or so, or even three years in advance. The company knows its yearly percentage of business, from which it is possible to forecast the probable percentage it will acquire in the next year or so, the third year being used only in a general way as a possible tendency. In the automotive industry it is even possible to carry this further by using the car registrations in each geographical division to forecast the probable sales of cars by makes or price classes in each of the nine geographical divisions used in the United States census chart. By adding the demand for any make of car in all the geographical divisions, it should check very closely with the demand determined from the production figures as outlined above.

With the potential demand determined, it is possible for a company to lay out its sales quotas—production schedules, purchases of materials, and financial arrangements and budgets—for each department for a year in advance and to adjust them each quarter according to the business cycle.

Manufacturers who employ this method or a similar one certainly will be in a much stronger position than manufacturers who continue to use the present inefficient methods. Facts show that no established company can acquire a great increase in its yearly volume except under the following conditions: (1) Introduction of a far superior product; (2) considerable reduction in price of an old established product.

A company may acquire a steady yearly increase in business by continually employing better methods than its competitors, or in other words, creating more favorable influences, both internal and external.

When scientific methods of forecasting future demands for products are generally employed in the basic industries, probably the variation in the business cycle will be considerably less. This would partially eliminate the sharp periods of inflation and depression and would tend to keep business on a more even basis than at present.

A STATISTICAL STUDY OF POOR RELIEF IN MASSACHUSETTS

By Katherine E. Howland, *Smith College*

Statistics relating to the nature and extent of poor relief have been compiled with far greater thoroughness in England than in America.[1] The central government of the United Kingdom gathers statistics of the number of paupers in institutions and also of the number in receipt of outdoor relief. Not only is there no central body of our federal government which collects statistics on poor relief, but in a large number of states not even figures are available for all the counties. In England and in Wales the total public expenditure for the support of the poor is also carefully ascertained annually, while in this country only a few of the states gather statistics relating to the amount expended from the public funds for the relief of the poor.[1] The field for historical investigation of poor relief in the United States is therefore limited to a few of the older states where records have been kept for a considerable period of years. It is the purpose of this study to discover if possible any tendencies relating to the nature and extent of poor relief in Massachusetts, where statistics relating thereto are available for a period of over fifty years, and incidentally to investigate the methods employed in recording the data bearing upon this subject.

Questions at once arise as to the percentage of the population of Massachusetts dependent upon public relief; the age, sex, and nationality of the recipients; the nature of the relief given; and the cost involved. Trends and analyses of the variability of these factors over a period of years may be expected to indicate significant aspects of the historical development of poor relief in Massachusetts. By attempting correlations of the percentage of the population in receipt of relief with such measures of business conditions as bank clearings or commercial failures, and average wages, we may still further investigate this problem of poverty.

The number per thousand of the population which received public relief during the years 1892 to 1920 is shown in Table I.[2] Vagrants, the insane, and the feebleminded are intended to be excluded from this enumeration, but owing to imperfect methods of classification and to differences in diagnosis this exclusion is incomplete. Previous to the

[1] Gillin, J. L., *Poverty and Dependency*, pp. 26–39. New York: The Century Co., 1921.

[2] Annual Reports of the State Board of Charity of Mass. 1875–1920. State Printers, Boston. This source was used throughout for all statistics of poor relief.

formation of the State Board of Insanity in 1898 the care of the insane was one of the main functions of the State Board of Charity, and therefore no great need was felt for differentiating the insane from the poor.

TABLE I

NUMBER OF POPULATION IN RECEIPT OF RELIEF CORRELATED WITH BUSINESS CONDITIONS

Year	Total no. in receipt of relief in Mass.	Number per 1,000 in receipt of relief	Number of cases of vagrancy	Number in receipt of out-relief	Ratio of out-relief to total relief	Indexes of wages per week in 17 industries in the U. S.	Per cent per firms in business in the U. S. of commercial failures
I	II	III	IV	V	VI	VII [1]	VIII [2]
1892	56,478	24.10	147,299	.....	.....	103.3	0.88
1893	54,819	22.88	141,481	.....	.....	102.2	1.28
1894	79,323	32.41	230,788	.....	.....	96.9	1.25
1895	82,544	33.01	304,244	.....	.....	97.8	1.09
1896	71,893	28.07	228,122	.....	.....	98.2	1.31
1897	78,122	29.79	241,230	.....	.....	98.2	1.26
1898	85,273	31.78	232,898	.....	.....	98.6	1.10
1899	85,591	32.28	207,326	.....	.....	101.3	0.81
1900	76,001	27.09	164,700	.....	.....	104.3	0.92
1901	75,913	26.68	170,986	.....	.....	107.2	0.90
1902	73,982	25.65	159,477	.....	.....	110.9	0.93
1903	72,094	24.65	129,997	.....	.....	115.5	0.94
1904	71,661	24.18	137,332	.....	.....	117.7	0.92
1905	81,413	27.10	117,513	.....	.....	120.3	0.85
1906	69,998	22.80	44,003	.....	.....	126.1	0.77
1907	64,310	20.50	26,224	.....	.....	130.7	0.82
1908	69,863	21.81	33,818	.....	.....	126.5	1.08
1909	78,923	24.13	32,798	.....	.....	125.8	0.80
1910	68,923	20.66	29,972	.....	.....	130.2	0.80
1911	71,441	20.97	32,300	.....	.....	132.3	0.81
1912	74,401	21.38	16,177	52,537	0.7061	135.7	0.98
1913	71,411	20.11	14,912	51,423	0.7061	.....	0.99
1914	86,425	23.86	16,190	65,059	0.7528	.....	1.10
1915	120,905	32.74	52,961	97,484	0.8063	.....	1.32
1916	118,012	31.68	31,662	95,935	0.8129	.....	0.99
1917	98,957	26.34	11,533	78,090	0.7892	.....	0.80
1918	89,582	23.64	11,366	72,936	0.8142	.....	0.58
1919	89,273	23.36	8,404	71,642	0.8025	.....	0.38
1920	75,868	19.69	3,082	62,770	0.8274	.....	.....

By Pearson's formula:

$r_{III, VII} = -0.833$ P.E. $= \pm 0.067$

$r_{III, VIII} = +0.426$ P.E. $= \pm 0.155$

Correlation after elimination of five-year moving average trend in each variable:

$r_{III, VII} = -0.622$ P.E. $= \pm 0.149$

$r_{III, VIII} = +0.438$ P.E. $= \pm 0.165$

[1] King, W. I., *The Wealth and Income of People of the United States.* London: The Macmillan Co., 1915. p. 203.

[2] Babson, R. W., *Half a Century of Business as Indicated by Babson's Business Barometers.* 1920.

Even had the attempt been made, knowledge of mental disease and defect has until recently been too meager to admit of a skilled diagnosis of insanity or of feeblemindedness. The exclusion of vagrants is almost equally difficult. Almshouses formerly served as winter quarters for many a tramp and as more temporary lodging houses for many more.[1]

[1] Kelso, R. W., *The History of Public Poor Relief in Massachusetts*, p. 113. Boston: Houghton Mifflin Co. 1922.

Vagrancy formerly presented such a serious problem to the administrators of poor relief that it is a subject worthy of special consideration. The number of cases of vagrancy for the years 1890 to 1920 are shown in Table I, column IV. In 1895 the cases of vagrancy reached the high-water mark of 304, 244, which should be compared with the cases in 1920 numbering only 3,082. The virtual disappearance of this problem is even indicated by Poole's *Periodical Index*. In the volume for 1900–1904 references are given to over 30 articles on the tramp problem, while in the index for 1918 the only reference to tramps is, "See Walking Trips." The word "tramp" does not even appear in the 1919 index.

Vagrancy has waned before the steady growth of public opinion resulting in the gradual development of more scientific methods of relief and of prevention. Formerly the general attitude was, "Why should we pay the tramp's board bills in jail? Tell him to move on." This universal evasion of the problem is well illustrated in a report from the Associated Charities of Lawrence, Kansas, which states that "the custom here with vagrants is (1) arrest; (2) fine; (3) as they have no money, they are put to work on the rock pile; (4) no guard; (5) the vagrants run off; (6) which is what was planned."[1]

Column VI, in the first table, showing the ratio of the number in receipt of outdoor relief to the total number relieved, probably represents a growing tendency away from institutional care; but the lack of comparable figures for indoor and outdoor relief has so narrowed the limits of comparison that the results do not warrant the determination of a general trend. It is worth noting, however, that the state has adopted a definite policy of outdoor relief for dependent children which is known all over the world as the Massachusetts plan of child placement. The state now has 7,000 minor wards, and of these wards all children who are placeable are located in foster family homes.[2]

Possible correlation of the number in receipt of relief with some measure of business prosperity is suggested by a comparison of Wesley C. Mitchell's [3] interpretation of business cycles with the rise and fall in the amount of poor relief as shown in Table I. It appears that a period of business prosperity is shortly followed by a fall in the amount of poor relief, while a business depression such as occurred in the years 1893–94, 1896, 1903–4, 1907–8, and 1913 is soon followed by a rise in the curve showing the number per thousand inhabitants who receive relief. The following statement, by Dr. Mitchell, of business cycles merits a close comparison with Table I, column III.

[1] *Atlantic Monthly*, vol. 101, pp. 744–53.
[2] Kelso, R. W., *op. cit.*, pp. 163, 186.
[3] *Business Cycles*, p. 88. University of California Press, Berkeley, 1913.

1892 General business prosperous; heavy gold exports.
1893 Panic; acute in May to October.
1894 Deep depression.
1895 Revival in the summer; followed by relapse.
1896 Return of depression; panicky conditions in financial markets.
1897 First half dull; second half, great improvement on crop situation.
1898 Prosperity rising rapidly.
1899 High tide of prosperity in general business.
1900 Slight pause in activity, followed by outburst of speculation in the autumn.
1901 Prosperity in general business.
1902 Prosperity continued.
1903 Financial liquidation; "rich man's panic"; general business still active.
1904 Mild industrial depression.
1905 Rapid gain in prosperity.
1906 High tide of prosperity.
1907 Crisis and severe panic in October and December.
1908 Deep depression in trade and industry.
1909 Recuperation; marked in second half of the year.
1910 Return of depression in second half of the year.
1911 Mild depression.

Of the correlations shown in Table I, that between the number in receipt of relief and index of wages in 17 principal industries in the United States computed by King (see explanation of source following the tabulation) shows the highest degree of association, there proving to be a negative correlation of 0.83 or, what is of greater significance, of 0.62 when the correlations are made of deviations around the trend lines. According to this association, variability in the number in the receipt of relief bears a distinct relation to the average wage paid. That is, there are a large number of persons who are not hopelessly and inherently paupers but who are dependent upon the marginal wage paid in industry. In times of prosperity these persons are independent. However, the wage which supports them is not only insecure, but it probably is also too meager to allow for saving against times of unemployment. The percentage of commercial failures appears to be less closely connected with the number of persons in receipt of relief, but the correlations found serve to substantiate the fact of a close connection between business prosperity and the extent of poor relief.

Let us now consider the age, sex, and nationality of the recipients of relief. Table II shows the age distribution of the recipients of relief for the year 1915 only. A close comparison, however, of the years from 1912 to 1920 reveals a lack of variability in the age distribution from year to year and shows that the mode of each year's distribution lies in the group from 5 to 9 years of age. Aside from any consideration of the age grouping of the general population, this fact is of great importance, showing as it does the need of a constructive program of poor

relief with the emphasis placed upon child welfare. Causes for the rise in the percentage of the total number in receipt of relief occurring in the age groups from 20 to 34 might well be sought in a study of immigration and of the conditions in modern industry where the productive capacity of the worker reaches its maximum at an early age.

TABLE II

AGE DISTRIBUTION OF RECIPIENTS OF RELIEF AND OF THE GENERAL POPULATION

Age	1915				
	Number in receipt of relief	Per cent of total number in receipt of relief	General population	Per cent of general population	Per cent of general population in receipt of relief
0– 5	13,524	11.19	363,297	9.84	3.72
5– 9	18,563	15.35	335,855	9.10	5.53
10–14	16,786	13.88	302,779	8.20	5.54
15–19	9,960	8.23	314,514	8.52	3.17
20–24	5,428	4.49	359,272	9.73	1.51
25–29	5,614	4.64	346,076	9.37	1.62
30–34	6,664	5.51	299,217	8.10	2.23
35–39	7,672	6.34	289,729	7.84	2.65
40–44	6,963	6.35	252,936	6.85	2.75
45–49	5,577	4.61	217,510	5.89	2.56
50–54	4,425	3.66	182,789	4.94	2.42
55–59	4,118	3.41	134,658	3.65	3.06
60–64	3,733	3.09	104,707	2.83	3.57
65–69	3,340	2.76	76,268	2.06	4.38
70–74	2,619	2.17	54,374	1.47	4.82
75–79	1,840	1.52	33,035	0.89	5.57
80–84	1,075	0.89	16,635	0.45	6.47
85–89	435	0.36	6,601	0.18	6.59
90–94	135	0.11	1,721	0.05	7.84
95–99	36	0.03	350	0.01	10.29
100–	10	0.01	46	0.001	21.79
Unknown	2,388	1.98	941	0.03	....
Total	120,905	100.00	3,693,310	100.000	....

Comparison of the age distribution of the recipients of relief for the year 1915 with that of the total population for the same year shows not only that the greatest frequency is found in the age group from 5 to 9 years and that there is a marked increase again from 20 to 44, but also that these maximum frequencies are out of proportion to the distribution of the total population. What are the chances that at a given age an inhabitant of Massachusetts will be a recipient of public relief? This question is answered in part by the last column of Table II. Five and one-half per cent of the general population in the age groups of 5 to 10 receive relief, as compared with 1½ per cent of persons from 20 to 24, and 21 per cent of those over 100 years of age.

The distribution of the recipients of relief according to sex from 1912 to 1920 shows that of the total number in receipt of relief the proportion of females has risen from 50 to 54 per cent. It is impossible to determine what has been the change in the general population during that period. According to the state census of 1915, 51 per cent of the total population were females; but of the total number in receipt of relief during the same year, only 50.23 per cent were females.

The recipients of relief during the years 1912 to 1920 are classified according to country of birth in Table III. It will be noted that from 1912 to 1916 a higher percentage of the foreign-born were in receipt of relief than of the native-born, after which period the conditions are reversed. This change may well be the result of the stoppage of immigration during the world war. Immigration to the United States reached the high figure of 1,218,480 in the year 1914, while the following year it dropped to 326,700 and continued steadily to decrease until 1918, when the total gross immigration numbered only 110,618.[1]

Russian-Polish farmers are popularly supposed to exceed by far the native population in thrift and efficiency. For this reason it is surprising to note the relative position of the Russian-Polish in this table, especially when it is taken into consideration that immigration from these countries is so comparatively recent. There is, therefore, a much smaller proportion of this nationality in the upper age groups than of the native population. It might prove enlightening to supplement this data with a study of Polish and Russian occupations in Massachusetts.

Obviously, the general form of each distribution in Table III is accounted for by the increase in the receipt of relief throughout the entire population during the war. Statistics as to the nationality of the recipients of relief are not available for a consecutive period of years before the war and it is therefore impossible to determine the normal trend for the different nationalities. At the close of the war a marked tendency to cluster more closely around the average will be noted. It is interesting to observe that the smallest rates of poor relief are found among those nationalities most sparsely represented in the total population. Perhaps they represent a more selected group than those nationalities which are more numerous.

The cost of poor relief is shown in Table IV. Although the program of poor relief is far more elaborate than it was in 1878, contrary to popular opinion, the cost to each inhabitant has not noticeably increased. The real effect of the war upon the cost of poor relief is obscured by the war inflation of currency. When this inflation is only

[1] Babson's Charts.

TABLE III

RECIPIENTS OF RELIEF CLASSIFIED BY COUNTRY OF BIRTH[1]

Birthplace	Number per thousand given relief								
	1912	1913	1914	1915	1916	1917	1918	1919	1920
Native	19.13	18.69	22.77	31.90	31.37	26.88	24.64	24.89	21.66
Foreign	22.39	20.78	24.79	31.37	33.73	26.60	23.41	22.59	18.07
Canada	22.53	22.43	24.33	32.98	31.03	17.23	22.69	17.00	16.01
England	23.15	22.24	24.91	31.93	29.83	24.28	19.94	17.45	13.62
Germany	13.24	13.97	14.08	19.00	18.15	14.84	11.71	11.64	8.33
Ireland	43.98	42.11	47.54	53.83	51.31	44.44	37.01	34.30	27.82
Italy	12.29	14.44	20.39	37.66	34.41	21.35	21.26	23.25	19.98
Russia	15.13	19.39	26.99	46.88	46.60	34.54	33.31	35.40	29.30
Scandinavia	8.91	8.51	9.69	13.38	13.12	9.43	11.25	10.62	7.82
Scotland	15.39	16.30	19.48	23.56	22.68	18.51	17.64	14.18	10.69
All others	15.22	8.90	12.79	23.70	24.36	28.07	16.78	20.92	13.51

[1] The distribution of the population by nationality could only be secured for the census year 1915, and as these figures were used in computing the percentage of the population according to nationality that received relief during the years 1912 to 1920, the results are subject to correction for every year except 1915.

roughly compensated for, as indicated in the table by the use of Babson's index of inflation, the rise in expenditure then appears similar to the increase in the amount of poor relief during the war, and of the same temporary character. It should, however, be noted, by comparison with Table I, column III, that the increase in cost during 1919 and 1920 is not accompanied by an increase in the receipt of poor relief.

TABLE IV

COST OF POOR RELIEF IN MASSACHUSETTS

Year	Total net cost [1]	Cost per inhabitant	Consumer's index of inflation [2]	Uninflated cost (consumer's)	Wholesaler's index of inflation [2]	Uninflated cost (wholesaler's)
I	II	III	IV	V	VI	VII
1878	$1,098,642	$0.787	....	....	...	....
1879	1,480,273	0.843	....	....	...	....
1880	1,412,388	0.793	....	....	...	....
1881	1,370,000	0.718	....	....	...	....
1882	1,426,702	0.772	....	....	...	....
1883	1,496,032	0.797	....	....	...	....
1884	1,542,824	0.808	....	....	...	....
1885	1,700,520	0.875	....	....	...	....
1886	1,720,795	0.891	....	....	...	....
1887	1,751,888	0.844	....	....	...	....
1888	1,826,872	0.849	....	....	...	....
1889	1,811,805	0.822	....	....	...	....
1890	1,878,710	0.839	....	....	...	....
1891	1,808,162	0.777	....	....	...	....
1892	1,869,244	0.789	....	....	...	....
1893	1,903,534	0.790	....	....	...	....
1894	1,993,263	0.814	....	....	...	....
1895	2,154,573	0.892	....	....	...	....
1896	2,190,515	0.858	....	....	...	....
1897	2,272,488	.867	....	....	...	....
1898	2,372,661	0.887	....	....	...	....
1899	2,522,410	0.919	....	....	...	....
1900	2,492,060	0.902	....	....	...	....
1901	2,551,074	0.898	....	....	...	....
1902	2,694,891	0.933	....	....	...	....
1903	2,780,751	0.952	....	....	...	....
1904	2,957,909	0.999	....	....	...	....
1905	3,096,035	1.033	....	....	...	....
1906	3,460,633	1.127	....	....	...	....
1907	3,240,763	1.034	....	....	...	....
1908	3,409,418	1.064	....	....	...	....
1909	3,712,689	1.135	....	....	...	....
1910	3,486,292	1.045	....	....	...	....
1911	....	....	....	....	...	....
1912	3,345,776	0.961	....	....	...	....
1913	2,846,085	0.801	....	....	...	....
1914	3,118,497	0.861	100	$0.861	100	$0.861
1915	4,047,471	1.096	100.5	1.090	101	1.085
1916	4,586,091	1.231	108.7	1.132	120	1.026
1917	4,663,940	1.241	131.7	0.942	187	0.664
1918	5,165,037	1.363	152.2	0.895	193	0.706
1919	5,846,121	1.504	172.2	0.873	219	0.687
1920	7,634,160	1.994	191.4	1.042	230	0.871

[1] Total net cost does not include cost of upkeep. Owing to the frequent obscurity in the distinction between state expenses and those of cities and towns, part of which are reimbursed by the state, the figures in this column are in the nature of a close estimate. For the year 1911, the state expenses are not given. By way of explanation it is stated in the annual report that "for the present report, comparable data of state care are not available."

[2] Babson's Reports.

Even this brief study reveals complexities in the analysis of poor relief. General business conditions, as measured roughly by bank clearings or commercial failures, and more especially by index of wages, affect the problem to an extent most difficult to determine. Age, sex, and nationality involve still other causes of variability and present other differences which should modify any program of relief. Study of poor relief from the standpoint of expenditure is complicated by the difficulty of securing a suitable index of depreciation of currency. Furthermore, questions as to the nature and effect of the relief given should not be considered apart from the growth and changes in public opinion, viewed as a main factor affecting the attitude of the recipients of relief as well as directing the public policy.

Accurate and complete records are essential to scientific study of such a problem as poor relief. Frequent changes in the method of securing and of recording data in the state reports at present make much valuable material incomparable. Some such plan as the following would render possible an increase in scientific knowledge which would be valuable in the administration of poor relief. The scheme should be drawn up by a federal board and adopted by every state that each may profit by the experiences of the others:[1]

I. Schemes for tabulation should be drawn up by a committee of expert statisticians to be followed under penalty of law for five years at least.

II. These schemes should cover all tabulations made by the towns and cities which enter into the state report, as well as those of the state, and also the method of final presentation of the combined material in the Annual State Reports.

III. At the close of the period prescribed by law, as recommended in I, the committee should incorporate useful suggestions into the scheme of tabulation to be followed during the next period. These changes should be clearly explained to all persons engaged in the compilation of statistics that are in the state report, as well as in the annual report itself, and the nature of the changes should be so stated that the data in any two consecutive periods may be rendered easily comparable.

IV. Special standard requirements needing greater emphasis than hitherto are as follows: (a) Outdoor relief should be carefully distinguished from indoor relief. (b) An analysis of new cases should be included as distinguished from the total number of cases. (c) A similar analysis of discharged cases should be made. (d) Classification of costs should be so standardized that the following totals for both the state and for the cities and towns are independent and exhaustive: (1)

[1] Such a scheme is of course impossible under the present legal organization of the United States.

Cost of persons relieved, not including expense of administration and upkeep. (2) Cost of administration and upkeep, the nature of all "incidental" expenses to be clearly specified. (3) Aggregate cost to the state, making due allowances for reimbursements, etc.

V. Summaries of the preceding years should include: (a) A brief comparison with the year directly preceding. (b) A more exhaustive review every 10 years. (c) An exhaustive statistical study every 25 years.

NOTES

THE EMPLOYMENT STATISTICS OF THE UNITED STATES EMPLOYMENT SERVICE

By Ralph G. Hurlin, *Russell Sage Foundation*

At the beginning of the year 1921 the United States Employment Service commenced the collection and publication of a monthly series of employment statistics as an adjunct to its other employment services. These statistics covered all important manufacturing industries and were obtained from a fixed number of firms in 65 large cities distributed throughout the United States. The data were published both in detail by cities and industry groups and in aggregate, the latter figures being interpreted as an index of employment fluctuation in manufacturing industries for the United States as a whole.

The scope and apparent nature of these figures were such as to raise hope that they would fill the need of such a representative index. The figures were made available shortly after their collection each month and they were much used during the depression period as indicating the course of industrial conditions. Criticism of the figures appeared, however, and in June, 1922, following the appearance of the report for May, their publication was discontinued by the Secretary of Labor with the explanation that this function was one to be reserved to the Bureau of Labor Statistics, a companion bureau in the Department of Labor.[1] In this case the question of whether or not statistics should be collected by the bureau requiring them in the exercise of its other functions or by another, primarily statistical, bureau, is clearly a matter of departmental administration and need not be raised here. The purpose of this article is rather to appraise the results of what has constituted a very extensive experiment in the collection of employment statistics.

In planning for the Employment Service statistics, economy of time in obtaining the data was considered an essential requirement, and in some degree this consideration determined the nature of the figures obtained. This and some other characteristics of the data make it of interest to see to what extent the figures appear to be a valid index of employment during the period for which they are available.

CHARACTER OF THE DATA

The index afforded by these figures is a large-city index, no data having been collected from establishments in rural communities or in

[1] Industrial Employment Information Bulletin, June, 1922.

small cities. It is also a large-establishment index. With one or two exceptions, only establishments employing under good conditions of business over 500 hands were included, and many of those included were excessively large establishments. Thus the total number of employees represented fluctuated around 1,500,000, which is 17 per cent of the total number of wage earners in manufacturing industries according to the census of manufactures for 1919, while the number of establishments was 1,428, or only half of one per cent of the total number of establishments enumerated in the census of 1919.

The figures themselves are described as pay-roll figures. They were obtained, in the largest district (including New York, New Jersey, and Pennsylvania) at least, upon printed schedules, bearing the single question of how many hands were employed on a specified date. These schedules were mailed to each establishment. Paid agents were placed in charge of the collection of the data in each of nine districts, but the information was actually obtained from the establishments in the several cities by unpaid agents, who secured the returns from the establishments and forwarded them to the district agents. The data were tabulated in the offices of the district agents, by cities and by industries, and in this form were telegraphed to the Washington office of the Service.

It is on the point of collection of the data that much criticism has been raised. Objection has been made to the fact that the collecting agents were not statistically trained and that in some districts due care was not exercised in securing authentic pay-roll figures. It is apparently true that there was variation from city to city in the precise method of collecting the figures, and in the case of at least one city the published figures themselves suggest careless or inaccurate returns. But this by no means characterizes the figures as a whole, and they furnish strong internal evidence of essential reliability.

COMPARISON WITH OTHER INDEXES

Examination of the statistics both by comparison with other indexes and by comparison of the trends recorded for the same industries in different cities suggests that the means of collection was for the most part sufficiently sound to give dependable results, and that the large-city, large-establishment index is representative. In the upper section of Diagram 1 the aggregate figures of the Employment Service are compared with those of the New York Industrial Commission for employment in New York State, for the 17 months from January, 1921, through May, 1922. The New York index is to less extent a large-establishment index, and it represents the entire state including the

small cities and other manufacturing communities. Because of the industrial composition of the state, this index has been regarded as a good indication of the course of employment in the United States as a whole. It is clear that the two curves reflect the same general tendencies.[1]

The lower section of the diagram presents the adjusted index of the Harvard Bureau of Economic Research for production in manufacturing industries in the United States. Because this curve has been

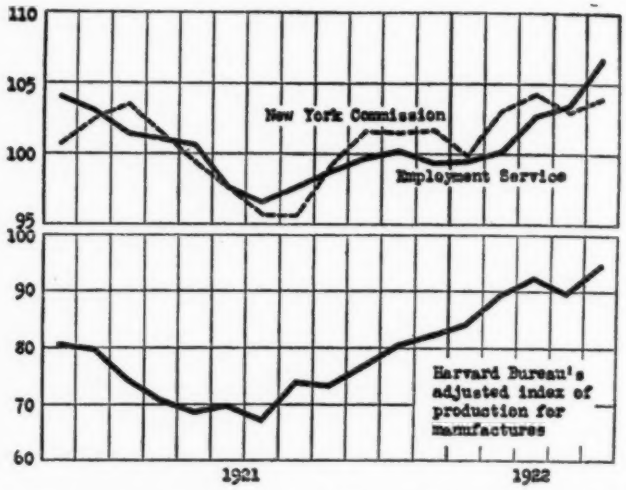


DIAGRAM 1. EMPLOYMENT INDEXES OF THE UNITED STATES EMPLOYMENT SERVICE AND THE NEW YORK INDUSTRIAL COMMISSION, AND ADJUSTED PRODUCTION INDEX OF THE HARVARD BUREAU OF ECONOMIC RESEARCH, JANUARY, 1921, TO MAY, 1922

"corrected" to show only cyclical fluctuation by eliminating other variations, it is not strictly comparable with the "uncorrected" curves above it. It is presented to show that it records the same turning point in the cycle, July, 1921, as does the Employment Service's employment index.

INDEXES FOR INDUSTRIAL AND GEOGRAPHICAL GROUPS

In order to test further the validity of the index, the figures have been broken up in each of the industrial groups into small geographical units. The industrial groups for which the figures are published are

[1] The December figure plotted in the Employment Service curve is not the published figure. The published figure was evidently too low, for a reason which is explained below. The figure here used represents the change registered in December by the aggregate figures minus those for the vehicle industry in the single city, Detroit.

the 14 groups used by the Census Bureau in the census of manufactures. For the present purpose the 65 cities were divided into eight geographically related groups, and curves were drawn for each industry in each of these geographical divisions. Conspicuous similarity was found between the curves for the same industrial groups with certain exceptions. Most of the exceptions were cases in which the curves represented too few employees to constitute a fair sample for the industry. In many cases such curves represented a single establishment or but two or three establishments in a given city. Other exceptions were cases which could be explained either by local industrial conditions, as in the case of the decline in the textile curve for New England due to the textile strike in that section in early 1922, or by the lack of homogeneity in the industrial group itself.

A portion of the composite chart of the curves for each of the 14 industries in each of the geographical divisions is shown in Diagram 2. It presents the curves for "all industries" and for three of the largest industrial groups, "iron and steel and their products," "textiles and their products," and "vehicles for land transportation." The first column of small diagrams is that for all industries. It shows striking differences in the curves for the several geographical divisions. These differences are clearly due to the influence of different industries. In Pennsylvania and Ohio the influence of iron and steel predominates. In New England the effect of iron and steel is balanced by that of textiles. In the North East Central division the vehicle group is sufficiently large to determine the character of the total curve.

The general form of the total iron and steel curve, shown at the bottom of the second column of diagrams, is seen in the iron and steel curve for each geographical division except that for the Pacific states. Here the figures are for few establishments and, in one case, for an establishment apparently undergoing rapid growth. This establishment would have had no effect upon the iron and steel curve for Pennsylvania and Ohio, but it has a controlling influence in the small sample for this industry in the Pacific states.

The textile industry shows similar striking agreement except in the case of the Pacific states, where again the sample is small. In this case the figures for one city, which account for the marked distortion of the curve in August, September, and October, 1921, are given throughout in round numbers, which fact suggests the possibility of inaccurate returns.[1]

The curves for vehicle industries illustrate lack of agreement because

[1] The published figures for the textile group in Los Angeles are: 510, 500, 500, 500, 525, 525, 525, 1,000, 1,000, 1,000, 425, 300, 350, 350, 350, 350, 455.

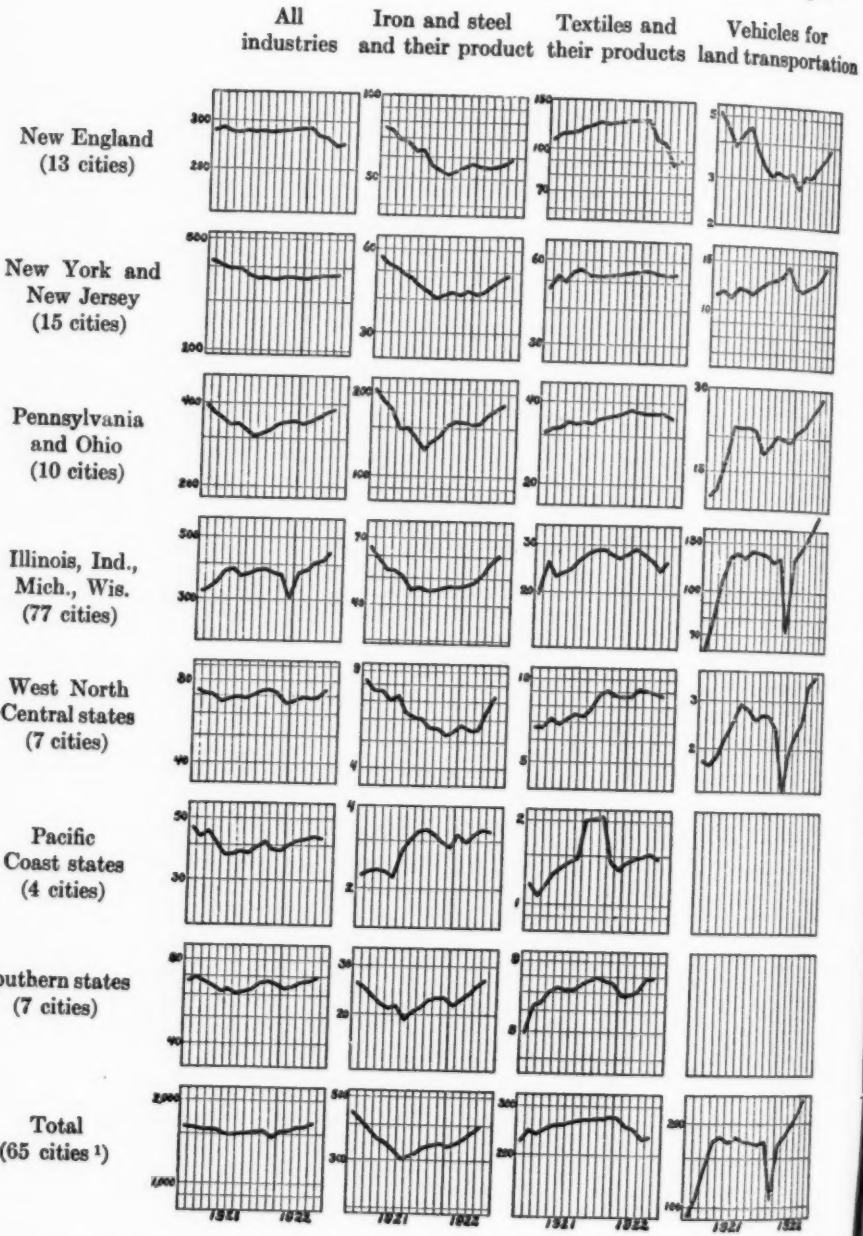

All industries
Iron and steel and their product
Textiles and their products
Vehicles for land transportation
New England (13 cities)
New York and New Jersey (15 cities)
Pennsylvania and Ohio (10 cities)
Illinois, Ind., Mich., Wis. (77 cities)
West North Central states (7 cities)
Pacific Coast states (4 cities)
Southern states (7 cities)
Total (65 cities [1])
1921
1922

DIAGRAM 2. EMPLOYMENT INDEXES OBTAINED FROM UNITED STATES EMPLOYMENT SERVICE DATA

Scale figures represent thousands of employees. Ratio backgrounds are used to permit direct comparisons between all curves

[1] Including two cities, Denver and Baltimore, not included in the sectional groups.

of difference in the composition of the group. The form of the total curve for vehicles is seen in the vehicle curves for the Pennsylvania and Ohio, North West Central, and North East Central groups, but the vehicle curves for New York and New Jersey and for New England are of different character. In the former case it is evident that the group represents primarily automobile manufacture, while in New York and New Jersey, and particularly in New England, it is probable that other vehicle industries are represented. No establishments in this group are included in the Southern or Pacific Coast cities.

So far as this form of analysis goes, it suggests that the figures obtained offer good indexes for the several industrial groups when these groups are similarly composed in the respective cities.[1] Other industrial groups within which good agreement is shown are: miscellaneous industries, which group probably here, as in the case of the census group itself, contains a large element of electrical machinery; leather; lumber; and paper and printing. In the food, chemicals, and stone-clay-and-glass groups varying employment trends are represented by component industries in each case. It is probable that for the purposes of an employment index changes in the census classification in the case of certain groups could profitably be made.

PERCENTAGE DISTRIBUTION OF WAGE-EARNERS IN EMPLOYMENT SERVICE INDEX AND IN ALL MANUFACTURES

	Employment Service index (average 12 months, 1921)	Census of Manufactures (average 12 months, 1919)	Difference
Textiles and their products	16.4	17.7	−1.3
Iron and steel and their products	22.5	17.4	5.1
Lumber and its manufacture	1.5	9.2	−7.7
Food and kindred products	8.4	7.5	.9
Railroad repair shops	3.9	5.7	−1.8
Paper and printing	3.2	5.6	−2.4
Vehicles for land transportation	9.9	5.5	4.4
Chemicals and allied products	4.8	4.7	.1
Leather and its finished products	3.2	3.9	− .7
Metals and metal products other than iron and steel	4.9	3.7	1.2
Stone, clay, and glass products	.9	3.3	−2.4
Tobacco manufactures	1.9	1.7	.2
Liquors and beverages	.1	.6	− .5
Miscellaneous industries	18.4	13.5	4.9
Total	100.0	100.0	.....

A fault of the index which appears to be of greater importance than the character of the figures themselves or the manner of their collection is the matter of industrial weighting. The accompanying table compares the industry weights in the employment index with the distribu-

[1] It should also be said that similar agreement is shown between these and other employment statistics for the same industrial groups, where similar classification of industries has been used.

tion of wage-earners shown in the census of manufactures for 1919. Iron and steel is excessively represented, as is the miscellaneous group, which resembles iron and steel to some extent. The vehicle group also has too great weight, whereas lumber is seriously underweighted. In view of the difficulty of controlling weights in a fixed-establishment index the differences shown in the table are not entirely surprising, however.

EFFECT OF THE FORD PLANT

It is much less certain that even approximately accurate weights are represented in the aggregate figures for individual cities. Here the disturbing effect of the excessively large establishments becomes evident. This may be illustrated by referring again to Diagram 2. In the case of vehicles in the North East Central section the curve represents about 130,000 employees, most of them contributed by one city, Detroit. A sharp drop is recorded in the vehicle figure for Detroit in December, 1921, which causes the deep depression in the vehicle curve for the section. It is probable that this was not an actual drop in employment but rather an interruption for inventory at the end of the year, since the loss was made up in the January figure. The amount of the drop for vehicles in Detroit was a little less than the number of employees in the Detroit Ford plant, and as practically all departments of this plant were closed for inventory at the end of December and as employment in other vehicle plants in Detroit is not sufficient to account for the drop, it appears that this establishment is accorded influence in the index sufficiently large to control the total curve for the North East Central section and that for the entire vehicle industry. It also accounts for the drop in this month in the aggregate curve for all industries in all cities, which causes the aggregate index to register its lowest point in December, rather than at the turning point of the depression in July. By eliminating the figures for vehicles in Detroit alone, a more representative figure for December is obtained. This was used in the construction of the curve in Diagram 1.

CONCLUSIONS

This analysis of the Employment Service figures is not assumed to be exhaustive. Nevertheless the following conclusions concerning these figures and concerning the larger problem of a representative index for manufacturing industries seem to be justified:

1. It is feasible to collect data for a representative employment index from a large number of cities in a short space of time and to make the figures available within a few days of the date of their collection.

2. The use of non-statistically trained agents is not necessarily a reflection upon the character of the figures.

3. A large-establishment and large-city index, if properly weighted industrially and geographically, is representative of employment trends for the country as a whole.

4. The geographical influence in the case of employment fluctuation is of slight importance compared with the industrial influence.

5. Although large establishments may be considered representative, care should be taken in the composition of an index to avoid excessively large establishments in cases where they alone control the index as a whole, or significant portions of it.

These statistics are no longer published in the bulletin of the Employment Service, but the machinery for their collection is still in operation in 65 cities and the figures are being returned each month for the use of the Employment Service. For the sake of completion of the record of an important period of employment fluctuation it is to be hoped that they may in time be made available to persons outside the Bureau. With minor changes in the list of establishments returning the figures and with only slight supervision of their collection, it appears probable that these statistics might be made to furnish a thoroughly reliable index of employment fluctuation for the country as a whole, and also reliable indexes both for particular industries and for individual cities. The latter service, which is of increasing importance, is one that is not now afforded by any other series of employment statistics.

THE DETERMINATION OF SECULAR TREND RECONSIDERED

By Holbrook Working, *University of Minnesota*

Statisticians will note with pleasure Professor W. L. Crum's discussion of the determination of secular trend in the issue of this Journal for June, 1922. It is one of the welcome signs of the growing interest of mathematicians in the problems with which scientists are struggling in the application of statistical methods in their various fields. The increasing use of statistical methods which economics is experiencing is common to many sciences. In all these sciences progressive workers are finding present knowledge of methods sadly inadequate. The necessity of devoting a large portion of their time to the acquirement of the requisite mathematical knowledge and the development of new and more adequate statistical methods has placed on such workers a heavy burden which is seriously retarding the progress of their work. The

prospect of aid from properly trained mathematicians is cause for rejoicing.

Professor Crum's discussion exhibits a weakness, however, to which the mathematician is naturally prone, and which seriously diminishes the value of his work from the point of view of the statistician. Proposing a method of calculating secular trend which is considerably more laborious than the present method, he neglects to prove that it introduces substantial advantages.

Professor Crum's point, it will be remembered, is that for many purposes the line of secular trend of a time series should be fitted by a method which gives the sum of the squares of the percentage deviations therefrom a minimum. The usual method, of course, gives the sum of the squares of the actual deviations a minimum. The validity of this point is unquestionable, and Professor Crum has made a distinct contribution in providing a method for obtaining the desired fit.

Unfortunately, for the practical statistician, Professor Crum has not gone on to present adequate evidence of the usefulness of his method. Approximations are always acceptable, even in theoretical mathematics, provided they are of a sufficient degree of closeness. It is the opinion of the present writer that the usual method of determining secular trend is such an adequate approximation. In support of this contention the following observations are submitted. It is hoped that they may lead Professor Crum or some other mathematician to provide a more adequate treatment of the subject as it concerns the practical statistician, to whom the labor of extended computations is a large consideration.

The method of *ratios*, to adopt Professor Crum's terminology, is claimed to give a line of secular trend which may differ notably from that obtained by the method of deviations, either in vertical position (a) or in slope (m) or in both.

The question may be attacked most simply by assuming a line fitted by the usual method, that of deviations, and considering the deviations —the d's—and the ratios—the $\frac{d}{x'}$'s—(x' being taken as the ordinate of secular trend corresponding to the deviation d). Consider first the values obtained for a. By definition we have $\Sigma d = 0$. The value of a for the line of secular trend fitted by the method of ratios will differ from that obtained by the method of deviations only in case $\Sigma\frac{d}{x'} = 0$. If the line of secular trend be horizontal, each value of d is divided by the same value of x' and it follows of necessity that if $\Sigma d = 0$, $\Sigma\frac{d}{x} \neq 0$. Hence, in this case a is the same for both methods.

If the line of secular trend be inclined from the horizontal, it is possible that the positive values of d might in general be associated with somewhat smaller values of x' than the negative values of d. In such a case we would have $\Sigma\frac{d}{x'}>0$, indicating that a would be larger if obtained by the method of ratios than if obtained by the method of deviations. Similar reasoning shows the possibility of having $\Sigma\frac{d}{x'x'}<0$ due to the positive values of d being associated in general with larger values of x' than the negative values of d.

In the absence of rigorous mathematical demonstration the above situations may be accepted as actual possibilities on the basis of the evidence submitted by Professor Crum. However, it is difficult to see how $\Sigma\frac{d}{x'}$ could be greater or less than zero by any considerable amount if the trend is approximately linear. Professor Crum cites three examples in which the values obtained are as follows:

Method	Value of a		
Method of deviations	2.9	9.9	398
Method of ratios	3.1	10.0	440

There is obviously either a miscalculation or a misprint in the third case as a value of 440 is impossible. The difference in the second case is not clearly significant, since a difference of one unit in the last place may be due to a difference of but one or two units in the succeeding place. The difference in the first case, assuming the computations to have been carried to a sufficient number of places, is significant, subject to two qualifications: (1) it may possibly be due to the fact that the method of ratios introduces an approximation;[1] (2) the series for which the values are computed is too short (seven observations) to be surely typical.

The writer has calculated the value of $\Sigma\frac{d}{x'}$ for seven series for which the data happen to be readily available. The series tested are the values of the annual production of hay, corn, oats, and wheat for 35 years, 1879–1913, as given by Professor Day (the first four of a group of twelve series published in one place,[2] and three series contained in unpublished material calculated by the present writer. In all but one of the seven cases $\Sigma\frac{d}{x'}=0$, within the limits of accuracy fixed by the

[1] *Cf.*, Jour. of the Amer. Stat. Ass'n, June, 1922, p. 213.
[2] *Cf. Rev. of Economic Statis.*, September, 1920, p. 258.

number of places carried in the computations. The exception occurs in the case of oats, where $\Sigma\frac{d}{x'} = -.056$, and $\Sigma\frac{d}{\frac{x'}{N}} = -.0016$. The latter is the more significant figure, as it is a measure of the vertical displacement of the line of secular trend. The values obtained for $\Sigma\frac{d}{\frac{x'}{N}}$ for the seven series are as follows:

Series	Years included	$\Sigma\frac{d}{\frac{x'}{N}}$
1 (Hay)	35	−0.0008
2 (Corn)	35	+0.0002
3 (Oats)	35	−0.0016
4 (Wheat)	35	+0.0004
5	23	−0.0001
6	22	+0.0002
7	20	+0.0002

In the case of oats, where the discrepancy is appreciable, the fit given by a straight line is not good; and an error of −0.0016 in its vertical position is insignificant.

The evidence from trial of specific cases, therefore, is distinctly in favor of the assumption that where the secular trend of a series is approximately linear, we shall have $\Sigma\frac{d}{x'} = 0$ (very nearly), and hence that the method of deviations will prove satisfactory in determining the vertical position of the line of secular trend.

Consider now the question of the slope of the line of secular trend. The method of deviations gives Σd^2 a minimum. The method of ratios should give different results only when a different line would give $\Sigma\left(\frac{d}{x'}\right)^2$ a minimum. Under what conditions will Σd^2 be a minimum, but $\Sigma\left(\frac{d}{x'}\right)^2$ be not a minimum?

The answer may be obtained from a consideration of Figure 1. The central group of points are those used by Professor Crum in his second example. The large black circles are at a distance from AB proportioned according to d^2 for the corresponding values of x. The large hollow circles are at a distance from AB proportioned according to $\left(\frac{d}{x'}\right)^2$ for the corresponding values of x. Adopting the concept of the

method of moments, these large circles may be looked upon as exerting a pull upon AB in proportion to their distance therefrom. If the solid circles be considered, these pulls are in equilibrium (Σd^2 is a minimum). But if the large hollow circles be considered, these pulls are not in equilibrium. The forces exerted about the lower end of the line have

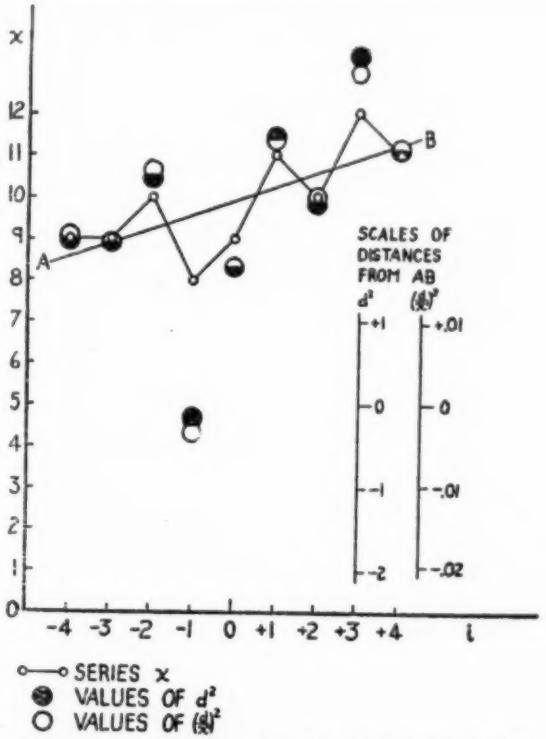


FIGURE 1. *AB* being placed to give the best fit for actual deviations, it will not in this case give the best fit for percentage deviations.

been increased (as compared with the solid circles), the forces about the upper end of the line have been decreased. It chances that the forces about the ends of the line are predominantly upward, so there is now a tendency to swing this line about its center. This is the explanation of the different slope obtained by the method of ratios.

Under what conditions will the forces at the ends of the line of secular trend be predominantly in one direction? Only when the trend is not linear. The curved line in Figure 2 is a parabola fitted by the method of least squares to the series shown in Figure 1. Clearly, the trend of the series is far from linear. Only the shortness of the

series prevented this curvature from being immediately obvious. Were the trend of the series truly linear, $\Sigma\left(\frac{d}{x'}\right)^2$ would be a minimum also, and the method of ratios would give the same slope to the line of secular trend as that obtained by the method of deviations.

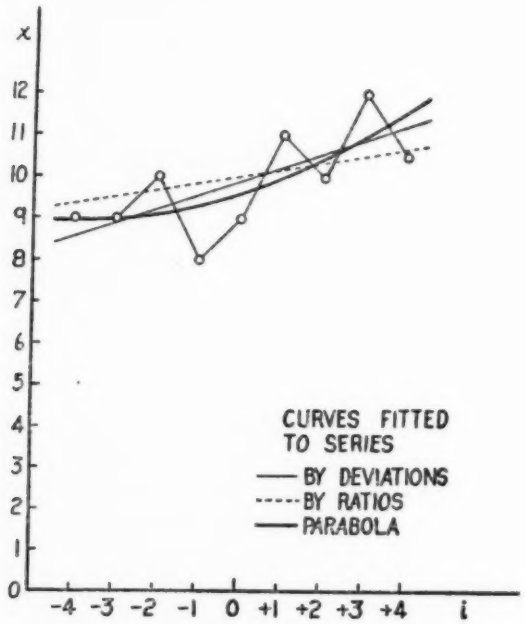


FIGURE 2. The curvature of the parabola shows that the trend of the series is not linear. This is the cause of the difference in slope of the two straight lines fitted to the series.

After a careful consideration of the principles involved, and a test in several specific cases, the following conclusion appears justified:

A straight line fitted by the method of least squares to a time series having an approximately linear trend gives a fit which is almost equally satisfactory whether actual deviations be considered or percentage deviations from the corresponding ordinates of secular trend. Professor Crum's criticism of the usual method is theoretically valid, but of little practical significance.

INDUSTRIAL CLASSES IN THE UNITED STATES IN 1920

By Alvin H. Hansen, *University of Minnesota*

In an article on "Industrial Class Alignments in the United States" published in the issue of this Journal for December, 1920, the writer made a study of the census occupational statistics for the years 1870–1910 for the purpose of classifying the gainfully employed into industrial classes. The purpose of this paper is to bring that study up to date.

The 1920 census material on occupational statistics differs somewhat in its classifications from the 1910 census. These differences, however, cause no significant difficulties so far as the present study is concerned. With a few unimportant exceptions the classifications described in the earlier study were followed, the only exception of any consequence being the transfer of chauffeurs from the "lower salaried" group to the "industrial wage-earning" class. This change seemed to be desirable because many persons formerly listed as "deliverymen" were classified in the 1920 census as "chauffeurs."[1] It is quite clear that the vast majority of persons returned as "chauffeurs" in 1920 were engaged in some form of industrial work.

The number of gainfully employed persons falling into each of the classifications described in the earlier study for each of the ten year periods from 1870–1920 is given in Table I. The percentage distribution is given in Table II.

A number of significant changes for the year 1920 will be noted. The number of farm laborers has enormously decreased, both absolutely and relatively. This is explained in the 1920 census bulletin on occupational statistics as follows:[2] First, the change in the census date from April 15 to January 1 tended to reduce the number of agricultural laborers for two reasons: (1) fewer laborers were employed in agriculture at that season and hence many laborers who would have been employed as farm laborers in April were returned in January under some other occupation; (2) children living on the home farm were in many cases not returned as either gainfully employed or attending school, whereas had the census data been for April instead of January they would doubtless have been returned as gainfully employed. In the second place, the war resulted in a considerable transfer of labor from the farms to the factories. For these reasons it can scarcely be doubted that the figure for agricultural laborers for 1920 is an under-

[1] Comparative Occupation Statistics, Fourteenth Census Bulletin, p. 3.
[2] Comparative Occupation Statistics, Census Bulletin, pp. 2–3.

TABLE I

	1870	1880	1890	1900	1910	1920
Farm laborers	2,885,996	3,323,876	3,004,061	4,410,877	6,143,998	4,178,637
Farmers	3,000,229	4,282,074	5,370,181	5,770,738	6,229,161	6,463,708
Proprietors and officials	581,378	807,049	1,347,329	1,811,715	2,879,023	3,168,418
Professional	414,708	666,338	1,114,507	1,565,686	2,074,792	2,760,190
Lower salaried	309,413	529,473	965,852	1,329,928	2,393,620	3,985,306
Servants	975,734	1,075,655	1,454,791	1,453,677	1,572,225	1,270,946
Industrial wage-earners	3,328,351	5,286,829	7,360,442	10,263,569	14,556,979	17,648,072
Unclassified	1,010,114	1,420,795	2,118,498	2,467,043	2,317,538	2,138,971
Total	12,505,923	17,392,099	22,735,661	29,073,233	38,167,336	41,614,248

TABLE II

	1870	1880	1890	1900	1910	1920
Farm laborers	23.1	19.1	13.2	15.2	16.1	10.0
Farmers	24.0	24.6	23.6	19.8	16.3	15.5
Proprietors and officials	4.6	4.6	5.9	6.2	7.5	7.6
Professional	3.3	3.8	4.9	5.4	5.4	6.6
Lower salaried	2.5	3.0	4.3	4.6	6.3	9.6
Servants	7.8	6.2	6.4	5.0	4.1	3.1
Industrial wage-earners	26.6	30.4	32.4	35.3	38.2	42.4
Unclassified	8.1	8.2	9.3	8.5	6.0	5.1

statement. On the other hand the 1910 figure for farm laborers is an overstatement compared to the figures given in the earlier census reports.[1] At all events it appears clear that the proportion of farm laborers to the total gainfully employed is not increasing and is probably no greater now than it was in 1890. This is striking, in view of the disappearance of free land and the increasing expense of setting up as independent farmers. The explanation no doubt lies in the increasing industrialization of the country with the consequent drain of labor to the cities.

Other significant changes appear in 1920. There was a considerable *absolute* decrease in the servant class and a very great *relative* decrease. This is in line with a steady decline since 1870. The servant class was two and a half times as important in proportion to other classes in 1870 as in 1920. This decrease is however more apparent than real. Personal service work has increasingly become industrialized. A large part of the work which was done in the home in 1870 was done outside of the home in 1920. Much house work has passed from the handicraft stage to the factory stage. The preparation of food is a case in point. Moreover, specialized occupations in the personal service field have been added from time to time with each succeeding census. For example, the 1920 census lists charwomen and cleaners, cleaners and renovators, porters, and elevator tenders in the personal service group —none of which appears in the 1870 census. It should be noted in addition that while servants have declined in importance the proportion of waiters has greatly increased, indicating a growing tendency to patronize hotels, cafes, and restaurants. Thus waiters increased from 107,044 in 1900 to 228,985 in 1920.

Other changes in 1920 are less marked and quite in line with the past development of the various groups. In the farming class there was a slight *absolute* increase, but a slight *relative* decline. The decline was not nearly so marked as in the 1900 and 1910 censuses. The professional class continued the steady increase it has shown since 1870. It is now exactly twice as important relatively as in 1870. Proprietors and officials, lower salaried, and industrial wage-earners—all subject to the influence of increasing industrialization—show increases quite similar to those previously registered in earlier censuses. The lower salaried class especially has been growing in importance, increasing nearly 40 per cent relatively to other classes from 1900 to 1910, and over 50 per cent from 1910 to 1920. This group is now nearly four times as important relatively as in 1870. The industrial wage-earning class increased relatively in much the same ratio that it has shown since

[1] Census of 1910, Vol. IV, pp. 26–29.

1870. The following table shows the increase for each period in the proportion of the total gainfully employed which were industrial wage-earners.

1870 to 1880	14.3	per cent	increase	
1880 " 1890	6.6	"	"	"
1890 " 1900	9.0	"	"	"
1900 " 1910	8.2	"	"	"
1910 " 1920	11.1	"	"	"

The "urban upper and middle" classes, including the proprietors and officials, professional and lower salaried groups, have increased from 19.2 per cent of the total gainfully employed in 1910 to 23.8 per cent in 1920. The "urban workers," including the industrial wage-earners and the servant groups, have increased from 42.3 per cent in 1910 to 45.5 per cent in 1920. On the other hand the "rural group," including farmers and farm laborers, has decreased from 32.4 per cent in 1910 to 25.5 per cent in 1920.

BRITISH BANKING STATISTICS

By W. F. Crick, *University of London*

One of the most striking characteristics of the scientific treatment of banking in the United States is the wealth of statistical material which is placed at the disposal of those professionally or academically interested in the subject. By contrast with this enormous stock of material the British banking statistics are woefully inadequate. The underlying explanation of this obvious contrast is of course the system of bank examination which has grown up throughout the period of American banking history, and the legal insistence upon the submission of periodical statements of financial condition. The National Bank Act and the Federal Reserve Act—to deal only with federal requirements—have given rise to an admirable body of evidence for use by the practical economist. The current reports supply an excellent basis for the interpretation of the present as an index of the future, without which the equipment of any far-seeing banker or business man is incomplete.

In Great Britain, however, with the exception of the statements of the Bank of England and the half-yearly statements of the other joint stock banks, there is no legal requirement whatsoever as to the publication of accounts. This fact is reflected in an excessive disposition to secrecy on the part of the British banker, which gives rise not only to an unwarrantable liability to public criticism, but also to insuper-

able obstacles to such detailed scientific analysis as is made possible in this country.

Herein lies a fundamental cleavage between the personnel of banking and its place in the economic structure of the United States and in that of Great Britain. The British banker, trained as a specialist from youth and very frequently the representative of a long line of banking ancestors, relies on tradition and instinct for much of the guidance which the American seeks in facts and figures. The American banker, often enough, is far from being solely a banker. The needs of the locality in which he works, and the impossibility of the thorough subdivision and specialization of functions that is a feature of the large American as well as the British banks, bring him into direct concern with the broader economic problems of his own community. Moreover, the "free" banking system cultivates an intense individualism and a self-reliance which demand a supply of something more tangible than tradition and instinct on which to base decisions.

In America, too, the conception of banking as a public function is much more widely diffused and recognized than it is in Great Britain. It is difficult to conceive of a member of the British Parliament attacking the "deadly deflation policy" of the Bank of England, and a general election fought on the issue of a single or double monetary standard can only be imagined as a fantastic illusion conjured up by a magician with a predilection for financial subject-matter. Public interest in banking as a vital part of the economic mechanism is in Great Britian practically non-existent. Hence there is no call for the publication of banking statistics, except the demand of distracted economists and students, such as would certainly arise in the United States should the present publicity be curtailed.

The *Federal Reserve Bulletin* and the reports of the Comptroller of the Currency, the Federal Reserve Board, and the State Bank Commissioners are incomparable for their wealth of detail and clarity of exposition. Moreover, their compilation by impartial agencies, and the irregular periodicity of the Comptroller's reports, makes "window-dressing" a more difficult operation than was the case under the old British system. Hence the problem for the investigator is not by any means that of amplifying or increasing the available data, but rather of interpreting and searching for their exact implications. The foreign observer, in particular, often fails to see the wood for the trees. The trees are admirably cultivated, but that only tends to make the foliage thicker. The problem in England, on the other hand, is to secure any statistics that will form a reasonably sound basis for considered deduction.

Accordingly, for the benefit of American business men and students who may profit from a skilfull study of the British bank statements from time to time, it may be of use to clear up some frequent causes of misconception, to fill in some of the pitfalls which lie in the path of the investigator, and to call attention to a comparatively new, and therefore almost unrecognized, source of banking information.

I

The weekly statement of the Bank of England is known and quoted everywhere as indicative of the state of health of the British banking organism. It was said at the outset that one of the two types of statement required by law was that applying to the Bank of England. When it is remembered, however, that this requirement was prescribed in 1844, under the Bank Charter Act, at a time when some of what are now the most elementary conceptions of banking theory had scarcely been formulated, it is not to be wondered at that this weekly statement is hopelessly antiquated. The Act prescribed that an account of notes issued, of the securities and gold and silver coin and bullion in the Issue Department, of the capital stock and deposits, and the money and securities in the Banking Department was to be published weekly in the *London Gazette*. The following is the form of statement used:

BANK OF ENGLAND

AN ACCOUNT PURSUANT TO THE ACT 7 AND 8 VICT. C. 32, FOR THE WEEK ENDING ON AUGUST 9, 1922

Issue Department

	£		£
Notes issued.....	144,029,625	Government debt.....	11,015,100
		Other securities.....	7,434,900
		Gold coin and bullion....	125,579,625
	£144,029,625		£144,029,625

Banking Department

	£		£
Proprietors' capital ...	14,553,000	Government securities ..	43,032,505
Rest.....	3,483,082	Other securities.....	75,450,390
Public deposits.....	16,238,919 [1]	Notes.....	18,487,415
Other deposits	104,500,212	Gold coin and bullion ...	1,820,255
Seven-day and other bills.....	15,352		
	£138,790,565		£138,790,565

August 10, 1922 Signature of Cashier.

[1] Including Exchequer, Savings Banks, Commissioners of National Debt, and Dividend Accounts.

The first statement, by itself, tells us nothing. The items "Government debt" and "Other securities" always total £18,450,000—the amount of notes permitted to be issued against securities. Even the term "Notes issued" does not by any means give ground for thinking that there are £144 millions of Bank notes in circulation. If the second statement be consulted it will be found that some £18½ millions are held in the Banking Department. Furthermore, by government accounts (which will be dealt with at a later stage), published independently, it is shown that over £21 millions are held in the reserve against currency notes. Hence, in fact, only £104½ millions are outside the Bank. It is still impossible to say how much of this amount is in actual circulation. A great part of it is held by the other banks as a part of their cash reserves, and no discrimination is made between notes and coin in the returns of these banks. These defections in the statistics are not obvious, and hence it is easy to jump at the conclusion that the public is making use of £144 millions of Bank of England notes.

Turning now to the second statement, the mysterious item "Rest" merely represents undivided profits, and hence, for all practical purposes, using the term in its accounting as opposed to its banking sense, reserve. This item simply serves as an indicator of the size of the next dividend, since the reserve is never allowed to fall below £3,000,000 and naturally increases as the time for a dividend declaration approaches. The deposit items are self-explanatory. Prior to 1877, however, they were much more instructive, for the item "Other deposits" was split up into bankers' balances and private deposits. These figures clearly gave a much better indication of the available credit power of the banks. Now they are practically worthless, for they fluctuate, not primarily through economic causes, but inversely in relation to each other. The reason is clear, in that the functions of the Bank as fiscal agent of the government bring about frequent transfers from one to the other, as, for instance, when income tax is paid, or the interest on government securities. Any extraordinary rise, such as that occasioned by war exigencies, does of course testify to an expansion of credit, but it is altogether too complicated a figure to be used as even an approximate index of the extent to which credit has been extended. The last item, notably small, but formerly of much greater magnitude, consists mainly of Bank Post Bills, a form of acceptance used for transmission purposes. A curious relic of the days of highway robbery in England, it now has the character, when accepted, of a bearer instrument.

The items on the assets side are more complicated. The word "securities" is used in a very broad sense, a sense altogether out of touch

with modern terminology. Thus "Government securites" includes not only long-time investments but Exchequer Bills, Treasury Bills, and even Ways and Means Advances to the government. Here are at least two fundamentally distinct classes of assets, hopelessly mixed. The next item is similarly constituted. It comprises not only long-term securities not issued by the government, but also bills discounted and all advances other than those extended to the government, quite apart from any question of the security behind them or their tenor. Nothing, therefore, can be gained from an examination of these items, except the fact that an increase in both *usually* betokens an expansion of credit, although it *may* quite as well mean increased holdings of long-term securities. From the point of view of the investigator the distinction made is not the one he would prefer. A separation of investments from bills and advances would be of much greater value than the present distinction between the persons liable on the instruments.

It is from these two items that the Bank's "ratio" is calculated. For all practical purposes the Bank note is a gold certificate. Hence the notes and coin held in the Banking Department are the cash reserves against the liabilities of the department. The liabilities against which they are held are simply the deposits. Hence the last two items on the assets side, divided into the sum of the third and fourth on the liabilities side, give what is perhaps the most important of the Bank's published figures, the "reserve ratio."

II

The condition of the joint stock banks is a much better criterion of credit conditions in the country than the statement just described, for, although the Bank of England is not so exclusively a "bankers' bank" as are the Federal Reserve Banks, yet it has comparatively little direct contact with the public. The outside banks, in contradistinction, are purely bankers for the public, and hence reflect the activities and sentiments of the public much more accurately, since directly, than does the Bank of England.

The joint stock banks, in addition to their half-yearly detailed balance sheets which, in common with all corporations, they are required by law to compile,[1] began a few years before the war to issue monthly statements. These were discontinued during the war, when they would have been invaluable, and have been resumed only since January, 1921.

[1] The Companies (Consolidation) Act, 1908, requires all public companies to file annually with the Registrar of Joint Stock Companies a "statement in the form of a balance sheet." These statements are almost invariably published, though it is not yet definitely established whether publication is a legal requirement. (See L. R. Dicksee: *Advanced Accounts*). A banking company, however, is specifically required to exhibit publicly, semi-annually, a statement of capital, liabilities, and assets.

With the limitations of date to which they are subject, these figures are very valuable. The banks differ among themselves, as would be expected, in matters of detailed presentation. Nevertheless, the separate statements permit of combination into tables which, though imperfect, are better than anything previously supplied. The composite figures[1] are compiled from the averages of the weekly returns which constitute the monthly reports of the individual banks. It is significant that these statements seem to have eluded the attention of American economic and financial periodicals which supply current information regarding conditions in foreign countries. The *Federal Reserve Bulletin* supplies a part of these data, but by unofficial publications they are practically ignored. Yet the figures are, in reality, much more significant than those of the Bank of England. The items on the liabilities side are quite straightforward. They include paid-up capital; reserve fund; notes in circulation; current, deposit, and other accounts; and acceptances. While it would be very valuable to have the current and deposit accounts separated, this combined figure is of much greater value than the figures of the Bank of England as a rough index of the volume of credit in existence. The assets comprise coin, bank and currency notes and balances with the Bank of England; balances with, and cheques in course of collection on, other banks in Great Britain and Ireland; money at call and short notice; investments; bills discounted; advances; liabilities of customers for acceptances; bank premises; and investments in affiliations. The first two items, of course, might be improved by further subdivision. The present condition of the Bank of England's ratio makes it very desirable to know how much of this most liquid of the banks' assets consists of coin, how much of notes, and how much of Bank balances. The second, again, is an unfortunate combination of two items that it might be desirable to separate.

It is difficult to make any very close comparison between these statistics and those published in America. The data relating to the Federal Reserve Banks are clearly out of the question, and the statements of the national banks are issued only five times yearly and include only one third of the total number of banks in the country. Probably the statements of the reporting member banks of the Federal Reserve System are the most directly comparable. The nine London clearing banks, with their branches all over the country, probably perform quite as large a proportion of the country's business as do the 800 reporting member banks in the United States. Moreover, the frequency of these reports, and the fact that they are compiled from returns supplied by the banks themselves, facilitate direct comparison.

[1] Published in the *Monthly Review* of the London Joint City and Midland Bank, and elsewhere.

For purposes of such comparison several of the items in the American figures would have to be thrown together, since subdivision is not carried so far in the British reports. Thus, "net demand deposits," "time deposits," and "government deposits" would have to be combined in order to compare with the British "current, deposit and other accounts." Again, "reserves with the Federal Reserve Banks" and "cash in vault" would have to be combined to correspond with the British "coin, bank and currency notes, and balances with the Bank of England." The comparison, however, falls short of completeness, owing to the omission of some items in the member banks' reports, such as "capital" and "bank premises," which appear in the British statistics. But in any case the comparison cannot be a close one.

For internal purposes, however, these statements are a distinctive and valuable contribution to banking data. Not the least of their valuable characteristics is the fact that they are now free from the charge of "window-dressing." Before these statements were issued it was almost universally recognized that the banks regularly contracted their advances and increased their Bank balances specifically in order to make a good showing in their half-yearly returns. Critics of the statistics are now disarmed, however, by the fact that the monthly figures are averages of weekly returns. It would clearly be impracticable for the banks to "dress their windows" four or five times a month. Hence it may fairly be argued that the returns as now published do not tend to exaggerate the liquidity of the banks' position.

III

In addition to these two sources of information there are two others of some interest and value. They are the clearing house returns and certain parts of the Government returns, issued weekly. The report of the London Clearing House is, naturally, the most important. The banking business of Great Britain is concentrated in London to a much greater extent than the banking business of the United States is concentrated in New York.[1] Hence the London returns furnish a more accurate index of the volume of business in England than the New York returns do of the volume of business in the United States. On the other hand, the London figures are affected so much by speculative activity that the elimination of this element cannot be effected as easily

[1] The returns of the London Clearing House, for example (as quoted by the *Economist*), showed a total of clearings for the year 1922 up to August 16 of £24,249,921,000, while those of the 11 leading provincial houses together in the same period reached only £1,269,843,600. By contrast, the New York clearings (according to the *Commercial and Financial Chronicle*) for the week ending August 19, 1922, totalled $3,027,100,000, which was less than the combined total for all other cities, viz., $3,587,486,621. Statistics of clearing operations of the Federal Reserve Banks and debits to individual accounts give additional support to this distinction.

as in the United States by using country clearings. It is true that the London Clearing House differentiates between "Town" and "Country" clearings, but even this discrimination is largely nullified by the fact that so much country business is financed through London that much of the provincial activity reflects itself, in monetary terms, in the Town clearings.

The Government Returns, so far as they relate to credit conditions, are useful in two ways. The Currency Notes Account gives the total of notes outstanding, and the gold, Bank of England notes, and securities behind them. The total of notes is a useful index of the volume of currency in circulation, though it is by no means exact, since some of the notes form part of the reserves of the other English and the Scotch and Irish banks. The Treasury Bills Account is useful indirectly. It shows not only the amount of bills offered and applied for during the past week, but also the average rate at which they were discounted. This is a fair index of the direction in which money is flowing, for, if the rate be £2 2*s* 0.9*d* per cent, as it was during the week ending August 11, 1922, it may be inferred that business is less active than when, as in the week ending August 5, 1921, they were selling at £4 10*s* 10*d* per cent. This figure, however, must not be too hastily relied upon, for there enters into the question the fact that the total of Treasury Bills offered has fallen very considerably, thus indicating that the demands of the government for this type of short-term money has decreased.

Besides these official figures there are various unofficial sources of financial materials for those interested in British economic conditions. The London Joint City and Midland Bank, for instance, compiles monthly figures of new capital issues, and produces the combined tables of the deposits, advances, etc., of the nine London clearing banks. The *Economist* tabulates the daily average prices of gold and silver. The *London Times* compiles weekly averages of the exchange rates on important financial centers. The British *Economic Review* issues an index of security prices. The London School of Economics (University of London), in addition to performing a similar tabulative and interpretative service to that undertaken by the Harvard Committee on Economic Research,[1] calculates a short money index from the averages of the Bank rate, bankers' deposit rates, three-months bills, and day-to-day rates.

Nevertheless, British statistics are notoriously inadequate. It would assist both business men and economists enormously in these days of increasingly scientific execution of the world's material affairs

[1] See the *Manchester Guardian Commercial*, Special Reconstruction Numbers. The work is under the supervision of Professor A. L. Bowley.

if the British banks would agree to issue statistics that would bear comparison with those now published by the member banks of the Federal Reserve System. It is confessedly a forlorn hope. Hence a proper understanding of the sparse information now available is especially essential to American and other foreign investigators.

THE PROBABLE ERROR OF THE DIFFERENCE BETWEEN TWO DIFFERENCES

By J. Arthur Harris

Given measures of two pairs of variables, w, x, and y, z, it is sometimes necessary to determine the probable error of the difference, d_1-d_2, where $d_1=x-w$, $d_2=z-y$.

A formula has doubtless been given somewhere in the extensive literature, but I do not recall having seen it. Since it must involve the value of $r_{d_1d_2}$ when there is any reason to suspect that the two differences may be correlated, a highly convenient method (which may be used when the correlations between the variables themselves are required and when variation and correlation are such that the ordinary theory of probable errors may be employed) is that based on the moments and product moments of the original variates, taken about 0 as origin. The formula for $r_{d_1d_2}$ has already been suggested [1] and employed in the study of growth increments.[2]

The moments $\Sigma(w)$, $\Sigma(w^2)$, $\Sigma(x)$, $\Sigma(x^2)$, . . ., and the product moments $\Sigma(wx)$, $\Sigma(wy)$, $\Sigma(wz)$, $\Sigma(xy)$, . . ., are generally required for other purposes of the investigation.

The constants for the differences are given by well-known formulas:

$$\bar{d}_1=\bar{x}-\bar{w},$$

$$\sigma_{d_1}^2=[\Sigma(w^2)+\Sigma(x^2)-2\Sigma(wx)]/N-\bar{d}_1^2,$$

and similarly for d_2, where the bars denote means and the sigmas denote standard deviations.

The product moment from which $r_{d_1d_2}$ may be deduced is given by

$$\Sigma(d_1d_2)=\Sigma(wy)+\Sigma(xz)-\Sigma(xy)-\Sigma(wz).$$

The standard deviation of d_1-d_2 may then be written

$$\sigma_{(d_1-d_2)}^2=\sigma_{d_1}^2+\sigma_{d_2}^2-2r_{d_1d_2}\sigma_{d_1}\sigma_{d_2}.$$

[1] Harris, J. Arthur, "Formulae for the determination of the correlation of size and of growth increments in the developing organisms." *Proc. Soc. Exp. Biol. Med.*, Vol. XVIII, pp. 4–5, 1921.

[2] Harris, J. Arthur, and Reed, H. S. "Inter-periodic correlation in the analysis of growth." *Biol. Bull.*, Vol. XL, pp. 243–258, 1921.

RECENT ACTIVITIES OF THE CENSUS BUREAU

June 30, 1922, marked the termination of the three-year decennial census period within which, under the census law, the reports of the Fourteenth Decennial Census were to be completed and published. By that date the manuscript and the statistical tables for all the reports had been sent to the printer, and all the fundamental statistics and many of the detail figures had been published as bulletins or advance press statements. There remained the reading and checking of the printer's proof for a number of large quarto volumes and the preparation of several monographs analyzing the statistics.

The Bureau has thus established a record in census taking, as it is safe to say that at no preceding census has the public been supplied with such an array of facts concerning population, social conditions, and industrial activities so promptly after the data were collected and within the period prescribed by law. The Bureau has issued seven series of state bulletins in connection with this census, two on the subject of population, one on agriculture, one on manufactures, one on mines and quarries, one on irrigation, and one on drainage. These series have all now been completed and published with the exception of the series relating to manufactures, for which a few state bulletins are still in the hands of the Government Printing Office. Of the eleven quarto volumes constituting the final reports of the Fourteenth Census, three have been printed, namely, Volume I, *Number and Distribution of Inhabitants*, giving the population of states, counties, incorporated places, and other minor civil divisions; Volume III, which comprises the state bulletins giving the population classified by sex, race, nativity, and age, together with statistics of illiteracy, school attendance, citizenship (naturalized or alien) of foreign-born, and dwellings and families; and Volume VII, which comprises the state bulletins for agriculture and is issued in three parts, one part for the Northern, one for the Southern, and the third for the Western states.

Of the projected series of interpretative studies or monographs, one, that on *Increase of Population* by William S. Rossiter, is in type and nearly ready for publication. The other monographs which are in various states of preparation include: *Recent Growth of Manufactures*, by E. E. Day, of Harvard University; *Recent Growth of the Mining Industries*, by G. A. Roush, of Lehigh University; *Immigrants and Their Children*, by Niles Carpenter, of Harvard University; *Illiteracy*, by Frank A. Ross, of Columbia University; *Farm Mortgages*, by Jesse E. Pope, of Washington; *Farm Tenancy*, by E. A. Goldenweiser, of Washington; *Wages in Manufacturing Industries*, by Paul F. Brissenden, of Columbia University; *Children in Occupations*, by Alba M. Edwards, of Washington; and *Women Housekeepers*, by Miss Bertha Nienburg, of Washington.

A large part of the force of the Census Bureau at the present time is engaged upon the first biennial census of manufactures, which relates to the year 1921 and is taken in accordance with the provisions of a clause in the Fourteenth Decennial Census Act which authorizes the Director of the Census to collect and publish statistics of the products of manufacturing industries every two years. The schedule is much briefer than the one used for the general census of manufactures

taken quinquennially, but it includes the number of wage-earners, the total wage payments, period of operation, cost of material, value of products, and in the case of many products, the quantity as well as the value. The results are being given out in the form of press summaries as rapidly as the work is completed, but as yet no printed report or bulletins have been issued.

In the meantime, of course, the Bureau is continuing its regular annual inquiries with regard to births, deaths, financial statistics of states and cities, and production and stocks of cotton, leather, skins and hides, and various other commodities.

The Bureau is making plans for the decennial census of wealth, public debt, and taxation, which will be taken next year and will cover the following subjects: Public indebtedness of the nation, states, and other civil divisions, 1923 and 1913; assessed valuation of property and amounts and rates of levy, 1860–1922 (for states, counties, municipalities, and other civil divisions); taxation and revenue systems of state and local governments; national and state revenues and expenditures, 1923 and 1913; county revenues and expenditures, 1923; municipal revenues and expenditures, 1923; and estimated valuation of national wealth, 1850-1922 (by states).

Preliminary work on this subject was begun last July, and it is expected that the section on taxation and revenue systems of state and local governments will be ready for publication before the close of the present fiscal year. Dr. Willford I. King, of the National Bureau of Economic Research, has consented to act as consulting expert for this inquiry, particularly with regard to the difficult problems of making an accurate estimate of national wealth. Other authorities will also be consulted regarding the scope of the inquiry and the methods to be followed.

Another important inquiry to be taken up next year is the census of defective, dependent, and delinquent classes. This includes in its scope the insane in hospitals, sentenced prisoners in prisons, jails, and reformatories, paupers in public almshouses, and feeble-minded in institutions for that class. It involves having a schedule filled out for each inmate present and each commitment or admission, giving age, sex, nativity, marital condition, and other data, including in the case of prisoners, offense and length of sentence, and in the case of the insane, the psychosis or type of mental disease. It should be added that the Bureau has not yet determined definitely the detail to be covered by these inquiries, but while the schedules to be used have not yet been formulated they probably will not differ very greatly from those used in the last preceding census of institutional classes.

In making the plans for the census of institutions and institutional classes the Bureau has held conferences with representatives of the National Committee for Mental Hygiene, the Child Welfare League of America, the American Institute of Criminal Law and Criminology, the National Probation Association, and other organizations and institutions.

A census of benevolent institutions is also to be taken next year. This covers orphanages and other institutions for children, homes for adults, hospitals, dispensaries, and institutions for blind and deaf. No individual schedules regard-

ing the inmates or admissions will be filled out except possibly in the case of institutions for children, but the report will give the institutional population on a given date, the number admitted during the year, the number discharged, and probably the receipts, payments, and value of property.

The Bureau is also planning to collect statistics of marriage and divorce relating to the year 1922. It is expected that this inquiry will be continued annually hereafter. As regards marriages, the Bureau heretofore has never done more than ascertain the number, but this time it is probable that a certain amount of detail will be compiled for those states where such detail is available in the state records. As regards divorces, the schedule will call for date of marriage; date of filing petition; who was libellant (whether husband or wife); whether case was contested; date of decree or judgment; number of years married; cause of divorce; kind of divorce (whether absolute or limited); number of children affected by decree; and whether alimony was granted. The following questions contained on the previous divorce schedules will be omitted this time: Date of separation; how notice was served (whether personally or by publication); whether intemperance was an indirect cause when not direct; and whether alimony was asked.

METHOD OF COLLECTING FINANCIAL STATISTICS OF CITIES AND STATES

Financial statistics of cities were first collected by the Bureau of Labor, under an Act of Congress approved July 1, 1898, which authorized the Commissioner of Labor "to compile and publish annually, as a part of the Bulletin of the Department of Labor, an abstract of the main features of the official statistics of cities of the United States having over 30,000 population."

The reports prepared by the Bureau of Labor for 1898, 1900, and 1901 covered a number of subjects, but always included the financial statistics; and since this work was taken over by the Bureau of the Census, in 1902, the financial statistics have been collected every year with the exception of 1914 and 1920, which years were unfortunately omitted on account of the press of the decennial census work. The statistics include the total and per capita receipts from revenues; total and per capita payments for expenses, interest, and outlays; total and per capita indebtedness; estimated true value and assessed valuation of property; and taxes levied, rates, methods of assessment, etc. Descriptions of accounting terminology and suggestions for uniform classification are also included. For a number of years reports were also published covering so-called physical statistics, such as statistics of police departments, fire departments, streets, charities, schools, and water, gas, and electric light systems, etc.

Financial statistics of states similar to those for cities were collected for the first time in 1915 by order of the Secretary of Commerce, and since then they have been collected annually, except in 1920.

From 1902 to 1921 the field work was performed entirely by agents of the Bureau of the Census who were sent to the cities and states. No assistance was asked of the city and state officials except that they permit access to their records

and explain any ambiguous entries in these records. The entire expense of the inquiry was borne by the Bureau of the Census.

These reports have proved to be of special assistance to the cities in the detection of antiquated methods and extravagant practices, and they are used extensively in the discussion of systems of taxation and improved methods of accounting. This being the case, it seemed only reasonable that the cities should bear part of the cost of the work. Accordingly, in 1921 the Bureau requested the city and state officials to coöperate to the extent of filling out the schedules, which work had heretofore been done by the agents of the Bureau. The Bureau was to compile and publish the reports as before. This new plan for securing the statistics was adopted at a time when agents of the Bureau were already in the field and considerable progress had been made in securing returns for the 1921 reports. The agents were directed to complete the reports on which they were then engaged, and to advise with the officials of other cities and states, rendering such assistance as might be necessary where the officials expressed a willingness to comply with the Bureau's request in future years but were unable to prepare the schedules required for the 1921 reports. The Bureau, when requesting the cooperation of officials, recognized the fact that a very considerable concession should be made in reducing the amount of work from that formerly required in the preparation of reports. Accordingly, new schedules, very much simpler than those formerly used, were prepared and mailed to the officials of the cities and states. As a result of this plan the Bureau obtained reports for 183 of the 253 cities having a population of over 30,000, and for 30 of the 48 states. Of the 183 cities included in the report, 33 were prepared by city officials, 21 from the annual reports of the cities and through correspondence with the officials, and the remainder by agents of the Bureau. However, the officials of only 14 of the 70 cities not included in the 1921 report refused to undertake to prepare the schedules; others first agreed to do so but failed to carry out their agreement, as no data were received; and in other cases where the work was undertaken the schedules sent into the office were in such condition that they could not be used even after the most extensive correspondence.

There were five states of the 18 not included whose officials refused to undertake to prepare the schedules. Reports for 27 of the states included in the 1921 report were prepared by the agents of the Bureau; reports for two states, California and New Jersey, were compiled by state officials; and for one state the data were taken from the printed reports, supplemented by information secured through correspondence with state officials. In almost every instance where the officials refused to prepare the reports they stated that they desired to coöperate but they did not feel that it was possible for them to prepare the reports properly, or they did not have sufficient clerical assistance. The net result of this plan was unsatisfactory.

In view of this experience the Census Advisory Committee recommended that the former method of collecting statistics be resumed, namely, the collection of data by agents of the Bureau of the Census, and that the statistics be placed on detailed schedules.

This matter was discussed at the meeting of the National Association of Comp-

trollers and Accounting Officers, composed principally of city officials. The discussion developed the fact that the members of the association and others interested would not have the same degree of confidence in reports prepared by state and city officials as in those compiled by a Federal agency, such as the Bureau of the Census. The contention was that city and state officials would be inclined to present the figures of their city or state in such a manner as to show the financial transactions of their respective governments in the most favorable light. Furthermore, it was thought that the several city officials could not prepare comparable reports. The Association accordingly adopted a resolution to the effect that the census reports on financial statistics of cities and states should be compiled by agents of the Bureau of the Census.

The Bureau of the Census, however, still feels that it is only reasonable that the various cities and states should bear a part of the expense of this undertaking which is costing the government about $100,000 annually, of which approximately $45,000 represents the cost of field work; and it is only fair to add that the city officials in 1921 very generally expressed a willingness to assist in any practicable way. Accordingly, it is proposed that in the future the Bureau send an agent to each city and state to be canvassed, with the distinct understanding that the city furnish clerical assistance to the agents of the Bureau, thereby reducing materially the expense of collecting the statistics in the field. Undoubtedly the National Association of Comptrollers and Accounting Officers can accomplish this by a persistent campaign among the officials in the different cities and states. Each state and city should designate a clerk familiar with their books to assist the representatives of the Bureau in compiling the census schedules. Upon this basis it is the intention to continue the work as recommended by the Census Advisory Committee, the National Association of Comptrollers and Accounting Officers, state and city officials, and others who are interested in public finance.

STATISTICS ON EMPLOYMENT: A REJOINDER

The idea in my footnote [1] at which Mr. Hornell Hart apparently took offense in his reply in the September JOURNAL [2] was not the result of any hasty or "careless" judgment on my part, but of long and careful consideration. I welcome the opportunity of stating more fully my objections to his method and of discussing his objections to mine—despite the unprofessional tone which he has chosen to use.

Giving attention to the four so-called "eccentricities" which Mr. Hart begins by pointing out:

(a) The first must, I think, be ignored, as Mr. Hart does not specify the incorrect footnote reference nor any of the misquotations, which of course I should have been glad to correct if proved.

(b) Mr. Hart's next objection is obscure. My point was that the typical behavior of the metal industry in reflecting the general industry cycle is even more

[1] This JOURNAL, Mar., 1922, p. 42, note 5.
[2] Pages 385–391.

strikingly demonstrated by employment than by production data. The coefficients in question were respectively 90 and 85 per cent. My statement as quoted by Mr. Hart is both correct and clear, so far as I can see.[1]

(c) Mr. Hart seems greatly to overrate the importance of always stating the probable error of correlation coefficients, because, whether the number of items be large or small, the probable error itself has a probable error, and moreover the very statement of a so-called probable error leads many persons to overlook the errors in the original data, which may be equally or more important. It is no oversight, as Mr. Hart implies, but a deliberate opinion on the part of many other writers as well as myself that the use of probable errors is very much overemphasized in time series of economic data.

(d) Mr. Hart is in error in expecting to find the mean of every cycle series 0.0 and its standard deviation 1.0. Indeed in many instances there would be ground for alarm if this were the case! The secular trend of a series is determined by the data embracing as nearly as possible some whole number of cycles. Once constructed, this trend may be extrapolated for short periods at either end of the base period for series which, like the ones under discussion, disclose little or no change of trend. This is a common and legitimate device, employed by Professor Persons and other leading students of time series for studying an index during parts of cycles adjacent to the original period. Therefore it ought not to seem disconcerting but perfectly natural to Mr. Hart that frequently $a \neq 0$ and $\sigma \neq 1$. Naturally, in computing correlation coefficients allowance is made for this fact by using the correction formula.

In appraising the numbered points to which Mr. Hart next proceeds, several considerations should be borne in mind:

Mr. Hart states in his discussion of point 2 that "correlation with business indexes cannot be considered an adequate criterion for judging the value of unemployment indexes"; and yet under point 1 he used such correlation of employment indexes with prices (65 *vs.* 52 per cent) as a condition sufficient "to remove any basis for supposing that Mr. Berridge's index deserves to be substituted," etc. These are clearly irreconcilable ideas.

In particular, price is an especially poor criterion for testing any index of employment because of the differences in their economic nature—one is predominantly physical while the other is not.[2] Professor Persons and Miss Coyle would be the last to claim that their ten-commodity index of prices measures *actual* industrial activity; prices are only an indirect measure. On economic grounds production is clearly much better suited as a test of the validity of employment indexes, though even here a higher correlation certainly does not *of itself* prove greater validity. Thus the employment index based on data for three states, which I showed in the June issue of this JOURNAL,[3] is more highly correlated

[1] The reader may judge for himself by referring to the original context, *Rev. of Economic Statis.*, prel. vol. 4, Jan., 1922, pp. 23, 24, including footnote 24. Compare Edmund E. Day, "An Index of the Physical Volume of Production," *Rev. of Economic Statis.*, prel. vol. 2, Dec., 1920, p. 367. As for the term "per cent," I see no essential difference between it and "hundredths," if this is Mr. Hart's point.

[2] For further discussion see *Rev. of Economic Statis.*, prel. vol. 4, pp. 24, 38–9.

[3] Page 236, Chart 4.

with pig iron production than is the New York index (89 *vs.* 85 per cent), but this fact is not enough to prove the superiority of the former without carefully considering the *a priori* grounds for accepting its components. For these reasons Mr. Hart's correlations of employment with prices are wholly inconclusive.

I cannot accept Mr. Hart's third argument, in favor of including all kinds of unemployment in the final index. It seems to me a cardinal tenet that unemployment due to depression, season, sickness, labor disputes, etc., forms not one problem but a series of very different problems, which should be studied separately as far as possible. True, I have not published any paper dealing with these types; but I have by no means "discarded" them. A final re-combination of all these factors is not only unnecessary but undesirable.

Practically all of Mr. Hart's fourth criticism seems to me beside the point, because it presupposes that I have been trying to measure the volume of unemployment, or at least the absolute numbers unemployed—an aim which I have disavowed repeatedly and emphatically. On the other hand, my purpose has been to develop as accurately as I could a picture of the *fluctuations* by using the index method.

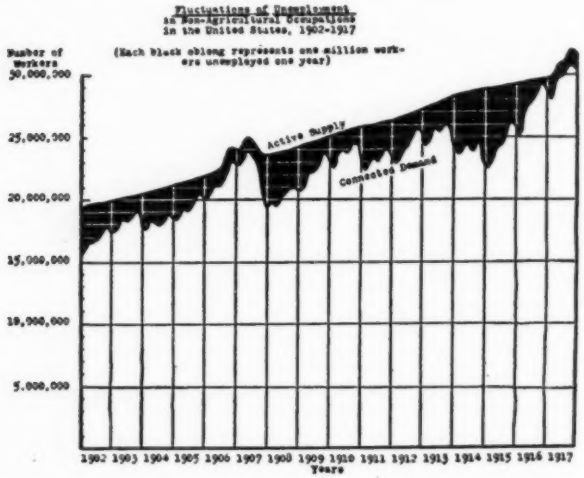
Under point 5, I agree that the samples which I have used in my index are not extremely large, though the three states comprise more than one fourth of the industrial activity of the entire country; but I deny that they are "biased."[1] I sought, for use in the index, only those data which seemed sure or at least likely to merit such treatment. Mr. Hart, on the other hand, accepted numerous data which I regard as untrustworthy for my purpose, and even more so for his; and other materials which are so fragmentary as to be safely used only as supporting evidence he has grafted directly into his measurements. Without rigorous selection of data it is impossible to obtain reliable results. As in my original reference to Mr. Hart's monograph, I concede its outward "comprehensiveness," and maintain that he paid dearly for this by accepting unsuitable materials. Comprehensiveness won on such terms is but an empty virtue.

Another important source of error, in addition to the uncritical selection of data, lies in the way Mr. Hart has attempted to use his data. One advantage of adopting the index method is that we thereby avoid the measurement of aggregates. Mr. Hart, however, has sought to measure (a) the aggregate volume of connected demand for labor, (b) the aggregate supply, and by subtracting the first from the second, (c) the aggregate volume of unemployment (see accompanying chart). In measuring connected demand, for instance, he attempted to translate data drawn from twenty-one distinct sources[2] into estimates of the absolute number of persons employed in each month from 1902 to 1917. A still further and perhaps even greater source of error lies in his determination of urban labor supply. This is based on annual estimates of population, birth and death rates, migration, etc., modified in an abrupt and somewhat arbitrary manner in 1906–7 and 1917 (see chart). I would go even further than Mr. Hart in pointing out variables that would have to be measured in order to obtain a base

[1] Mr. Hart greatly exaggerates the differences among industries as to the normal timing of employment cycles. For evidence on the similarity among localities see this JOURNAL for June, page 232.

[2] Enumerated on p. 49 of his monograph, from which the chart is also taken.

line suitable for the purpose he has in mind; indeed they are so numerous, and so imperfectly represented in available statistics, that I regard the problem as virtually an indeterminate one. If the secular trend is, as Mr. Hart claims, a "crude and unsatisfactory" datum for measuring the fluctuations of an employment index, then his supply line is still more "crude and unsatisfactory."



Finally, the method of obtaining a residuum by subtracting one of these large and unreliable absolutes from the other magnifies the seriousness of the errors in them both; this seems self-evident.

In view of all these considerations, I am unable to convince myself that my own methods or results have been affected at any vital point by Mr. Hart's attack, or that I should change my original verdict on his research.

W. A. BERRIDGE

A PROPOSAL FOR A STANDARD SERIES OF STATISTICAL REFERENCE BOOKS

For many years there has been a feeling among trained and experienced statisticians that standards of professional skill should be established in the statistical field comparable in a general way with those which are recognized in many other professions. Attempts have been made at various times to formulate such standards, but these attempts have resulted in little more than an increased realization of the difficulty of outlining professional requirements in a scientific field as broad and complex as that of statistics.

During the early part of the present year this question was brought up for consideration at a meeting of the Board of Directors of the Statistical Association. After considerable discussion it was agreed that the real need at present was not

that of establishing a basis of certification for statisticians of varying degrees of professional skill, but was the more fundamental one of securing an agreement among academic and practicing statisticians as to the cultural background and the elements of descriptive and analytical statistical technique which should form the basis for a general statistical training. As the result of this conclusion a special committee [1] was appointed to consider the matter in further detail and to make its recommendations to the Association.

In approaching this very difficult assignment the committee's first effort was to determine the specific form that final action should take to be of real service to the statistical profession. It was obvious that a mere statement, in general terms, of the fundamental importance of cultural studies and of certain elements of statistical technique would be of little practical value at the present time. The immediate need of teachers, students, and practicing statisticians appeared to be for an authoritative and rather detailed outline of principles and practice in the elementary and intermediate rather than the advanced ranges of statistical science.

With the problem thus newly defined, the committee then undertook to survey in a general way the existing statistical treatises and textbooks in English. This survey indicated the existence of many excellent elementary and advanced texts, and of many valuable treatises on special statistical applications. There appeared, however, to be a distinct lack of what might be called intermediate texts, representing the dividing line between the skill of the statistical specialist and that knowledge of statistical principles and practice which should be the common property of all well-trained statisticians.

As a result of this survey, the idea developed that the Statistical Association might very properly undertake to fill the existing gap by securing the preparation of a coherent and properly related series of statistical reference books. A further study of the question indicated rather clearly that this series could not be expected to replace textbooks prepared for purposes of instruction in colleges and universities, nor could it be expected to replace advanced treatises on statistical methods and statistical applications. It seemed clear that the proposed texts should be designed primarily for reference purposes, and that it would be particularly necessary to have them prepared under such supervision as would make them represent in the subject-matter as a whole, and in the relative space assigned to the various elements of illustrative subject-matter, a sound and authoritative consensus of up-to-date statistical opinion and experience.

In continuation of its study of this problem the committee has developed a definite proposal which it is the purpose of this article to submit to the membership of the Association, in advance of the annual meeting, for their comments, criticisms, and constructive suggestions.

The proposal, in brief, is that there shall be appointed, at or following the annual meeting, a general committee on statistical reference works with a membership properly balanced between academic and practicing statisticians, and having power:

(a) To appoint an editor-in-chief and special editorial board, this board, ex-

[1] Consisting of I. I. Dublin and M. C. Rorty.

cept for the editor-in-chief, to be made up of the several writers who undertake to prepare individual texts.

(b) To appoint an advisory board, which shall coöperate with the editorial board in the preparation of the required texts.

(c) To select a publisher and determine an equitable basis for the apportionment of royalties as between and among the several writers and the editor-in-chief.

(d) To make initial recommendations to the editorial board as to the general scope and plan of the series of reference volumes.

It is the feeling of the present committee that the general committee should serve simply to organize the project in its initial stages and to select the permanent working organization as above indicated.

Consultation with various members of the Association indicates that the proposed series might properly include the following volumes:

1. Outline of statistical principles and methods
2. Statistical sources and the collection and tabulation of statistical data
3. Graphical presentation of statistical data
4. Graphical analysis of statistical data
5. Statistical mathematics. Theory of probabilities. Theory of sampling. Statistical accuracy and error
6. Frequency distributions and averages
7. Curve fitting and correlation
8. Statistical forecasting and the analysis of time series
9. Index numbers and other statistical indexes

The preceding titles are of course very tentative, and indicate only in a rough way the possible arrangement and scope of the several volumes.

It has further been suggested that the individual volumes be restricted, so far as practicable, to not more than 200 pages each, in order that the selling price of the series may be kept within reasonable limits. These and other details of the series may very properly be left for consideration by the proposed general committee and editorial board. However, it is the suggestion of the present committee that the initial volume, "Outline of statistical principles and methods," should consist, in the main, of a series of chapters directly related to the remaining volumes of the series. It is assumed in this connection that the writer of the initial volume will be in position to coöperate directly as to each chapter with the writer of the related volume in order to secure the necessary continuity of treatment.

The present thought is for the initial volume to include the range of descriptive and analytical operations that should be covered in elementary statistical courses for engineers, economists, sociologists, biologists, medical students, etc., as well as for those who expect later to specialize in statistical work. Also, in the office of the practicing statistician it is felt that the elementary text should serve as the primary reference work for "junior" statisticians and that the series as a whole should adequately cover the field for "statisticians" and "senior statisticians," leaving to advanced texts the field covered by statistical specialists.

For the purposes of the annual meeting, at which the plan will form the basis of

a round table discussion, it will be helpful if members who are interested will mail promptly to Professor Robert E. Chaddock, American Statistical Association, Kent Hall, Columbia University, New York City, their comments and suggestions as to the proposal as a whole and in detail.

It should be noted particularly, in this connection, that it is possible for the series to be issued either under the auspices of the Association or as an independent venture to which the Association will not be directly committed. A notable precedent for direct action by the Association may be found in the work of the British Institute of Chartered Accountants in connection with actuarial science. On the other hand, there may be some members who would endorse the project as an independent venture, but who would be unwilling to see the Association assume direct responsibility.

A COMMUNICATION

In the September number of this JOURNAL there appeared a review[1] of *Costs, Merchandising Practices, Advertising and Sales in the Retail Distribution of Clothing*, published by the Bureau of Business Research of the Northwestern University School of Commerce. Because the reviewer has so fully misunderstood the purposes for which this study was undertaken and the methods of realizing them, I am asking the Editor for permission to insert this brief note.

When the Bureau of Business Research undertook this study, it was done

> . . . with the distinct feeling that it would not be worth while, either for the Bureau or for the trade, to make a superficial investigation. Accordingly, the questionnaire . . . was comprehensive in outline and detail in the various points covered. . . . In view of the fact that the details were submitted, it would not be in keeping with the intent of the Bureau of Business Research, nor fair to those who furnished the information, to present the facts in a general and condensed manner.[2]

With this idea in mind, it was thought

> . . . not to be sufficient to group roughly all stores into a single class and to express costs and merchandising practices in the form of a single average or total. Averages, unsupported by the detail which make them up, or without reservations as to the diverse conditions which obtain, may be not only confusing, but also misleading. By giving more than averages, that is, by classifying stores into groups, according to size and location, by distinguishing them by cities and by positions within cities, and by treating the years 1919, 1918, and 1914 separately, each merchant, who has sufficient interest, may compare his own store with others similarly placed and judge of his own activities and costs in terms of those prevailing for the group to which he belongs.[2]

Other statements relating to the purpose of the study occur at various places throughout the different volumes.

The aim or purpose of undertaking this study was realized by reducing to a comparable basis the different conditions which seemed to be pertinent concerning each of the topics discussed. Accordingly, stores are fully classified by size and location, by position within cities, by types of buildings occupied, by floor space utilized, by types of goods sold, etc., the purpose throughout the different

[1] By Mr. H. K. Herwitz.
[2] Page 8.

volumes being to classify the stores so that the averages would have as nearly as possible a clear and unmistakable significance. In other words, it was felt that only like things could be compared, and a definite attempt was made to group the stores with respect to conditions which were similar. Had the Bureau been content with crude generalizations, it would not have utilized more than 600 pages in which to describe the expenses and merchandising practices of retail clothing stores, nor would it have taken the pains in summarizing its results to call attention to the fact that the general figures which are given are significant only to the extent that they accurately characterize the details of which they are made up. It may be worth while, in this connection, to cite a typical expression which is found at various places throughout the study. In summarizing the contents of Volume 1, these words are used:

> It (the summary) should be considered as a summary and read in connection with the corresponding parts in subsequent pages. To summarize briefly such a study as this requires broad generalizations and does not admit of all the conditions and limitations affecting the conclusions being enumerated. It can almost be taken to be a self-evident truth that general statements are inexact. It is hoped that the reader will not content himself with a hurried perusal of the summary. The really significant part of the discussion is not in the summary, but in the details which support it.[1]

Realizing the limitations of any brief summary as an accurate characterization of the different parts of the study, the different volumes[2] are provided with summaries at the close of each independent part. In these summaries, relating as they do to specific items of expense or trade practices, attention is frequently called to the variations in the averages for stores of different size, location, merchandising practices, etc., and to the fact that the averages do not reveal all of the peculiarities reflected in the data which they include. The reader repeatedly is urged to consult the various tables in the discussion and not to rely upon the summaries alone. Further evidence that the Bureau was unwilling to summarize the data by the use of averages of non-homogeneous groups is found in that part of Volume VI which relates to expense- and trade-tendencies. This summary features the different conditions which obtain for stores of different size, location, position, etc.

The reviewer of the study objects to the use of averages, as distinct from modes or medians, to characterize the conditions obtaining in the different store-groups. In so doing, he seems clearly to have misunderstood, or at least not to have appreciated, the purposes of the study and the methods of realizing them. Briefly, these purposes when put under two heads were: first, to establish the "absolute and relative amount of sales, rent, wages and salaries, advertising, etc., for stores of different size, location, age, merchandising and accounting methods," and second to determine "the tendencies of the amounts to decrease, increase, or remain constant as stores increase or decrease in size or change in location and operating conditions."[3] In order to realize these purposes, the arithmetic mean is the only simple measure that can be used. In many instances modes are indeterminate and medians of no significance. Some single precise unit had to be

[1] Page 10. Similar statements are found on pages 103, 185, 269, 277, 388–89, 395, 496, 499, 571–72.
[2] Except Vol. I.
[3] Page 571.

adopted. Modes would not have sufficed. Frequently there are no clearly defined modes. Indeed, from many points of view, the tendencies for amounts of expense, etc., to increase or decrease under different conditions of store operation are far more important than the actual amounts of expense at a particular time as revealed by any measure. Such tendencies can be measured satisfactorily only by the arithmetic mean. Time and again throughout the study, as noted above, merchants are not only advised but entreated not to compare their own expenses with those of all stores, but by using the index and table of contents to select those particular store-groups, those particular locations, sites, and operating conditions which most nearly correspond to those obtaining for their own stores, and even then to look upon the amounts of expense, etc., as only generally indicative of the conditions obtaining at the time or place discussed.[1]

The reviewer has compiled a table, for the different items of expense for *all* of the stores treated as a unit, wherein the averages differ markedly from modes and medians. To have prepared such a table rather than one in which the respective averages are computed for stores of similar size, location, position, etc., is to do precisely the thing against which the Bureau so frequently and painstakingly protested; that is, to take general averages in preference to those which represent the greatest detail obtainable for a particular situation. Moreover, the table compiled by the reviewer forcibly illustrates the difficulties which would have been encountered in comparing the stores in the detail which was required had the Bureau chosen to use modes or medians. In many cases modes or medians derived from the totals, even under the detail in which the stores are tabulated, would not have made it possible to have compared stores group by group. To base comparisons, as the reviewer does, on modes and medians for *totals* alone, as contrasted with those for the detailed groups, is inexcusable, because to do so involves averaging things which are patently non-homogeneous.

The study is also criticized because relatively few frequency distributions are given. The Bureau would like to have printed the frequency distributions not only for the totals but for every item of the detail. This, however, seemed unnecessary, and quite impossible within the amount of space available. Accordingly, it chose the method of classifying the stores in great detail and then of computing arithmetic means. These are precise and allow comparisons to be made from store-group to store-group. It did, however, supply the frequencies for the different elements of expense, for sales, turnover, depreciation, discounts, fixtures, outstanding accounts, returned goods, sales per square foot of floor space, etc., for 1919, or for another appropriate year, and called attention repeatedly to the fact that averages should be interpreted in terms of the number of instances used and in conformity with the peculiarities exhibited in the frequency tables.

Attention is also called to the fact that more identical stores are shown for the three years 1914, 1918, and 1919 than for 1914 alone. The reason for this would have been patent to the reviewer had he observed the plan under which the data are presented. For each of the three years for the identical stores it was neces-

[1] The discussion on pages 167, 192–193, 203–204, 215, 263, 314–315, 388–389, 467, 469–470, 478, 523, 533–34—selecting only a few places where attention is given to this topic—is sufficient to indicate the care that is used to caution the reader not to use averages without the supporting data to which they apply.

sary, as the data are presented, for each store to report only sales and rent, while for the single year, 1914, it was necessary that sales, rent, and total expense be reported. The use of a different number of stores in the various presentations is typical of the method followed throughout the study. The reason for doing this is referred to by the reviewer in another connection and should have been evident to him in this case.

In brief, the reviewer, while describing in a general way the contents of the Bureau's study, forgets the purpose of the study and takes little note of the methods which have been used to realize it. Because of the fact that statistical data are always susceptible to misuse, that they may appear to be exact and at the same time be full of error, and that many people are credulous to a fault in the use of them, the Bureau took special precaution to guard against its findings being improperly used. If its efforts in this respect have been almost completely overlooked by the reviewer, it can only hope that such will not be the case on the part of others into whose hands the volumes come.

HORACE SECRIST

The Bureau of Business Research,
Northwestern University, School of Commerce

A REPLY

BY H. K. HERWITZ

On the essential points Professor Secrist and I do not differ. The arithmetic mean is easy to find and easy to compute. Modes and medians are not. As Professor Secrist knows, however, and as the illustrations that I take from his study clearly show, the arithmetic mean, although it is easy to calculate or even because it is easy to compute, frequently does not show what it professes to show. This is my point. And it still seems to me true that matters of this kind should not be left to the scrutiny of ordinary readers of books of this sort, who usually are so little practiced in the art of statistics.

With regard to Professor Secrist's second point that the "detailed" tables would have yielded different results than the "total" tables, I can only add that an examination of the "detailed" tables does not confirm Professor Secrist's statement.

MISCELLANEOUS NOTES

Vital Statistics Courses. The following information on vital statistics education in schools of public health was collected by the Committee on Vital Statistics of the American Statistical Association:

University of Chicago. Department of Hygiene and Bacteriology
Vital statistics and epidemiology. Professor E. O. Jordan.

Harvard University. School of Public Health.
An elementary course of lectures and laboratory work designed to familiarize the student with descriptive vital statistics, with graphic representation and with the basic theory of probabilities and correlation. Edwin B. Wilson, Professor of Vital Statistics.

Johns Hopkins University. School of Hygiene and Public Health.

Several courses designed to meet the needs of persons training for public health work and for students intending to specialize in statistical work as teachers, investigators, or administrators.

(a) Introduction to Vital Statistics
Methods of collecting, tabulating, adjusting and interpreting statistical data regarding human life, and review of chief statistical results of the past in this field.

(b) Laboratory Work in Vital Statistics

(c) Advanced Statistical Theory
Methods of statistical analysis. Curve fitting. Graduation of data. Life tables. Correlation.

(d) Colloquium on Practical Vital Statistics
Training in critical application of vital statistics to specific problems in biology, medicine and hygiene.

(e) Statistical Measurement of Public Health Activities

(f) Investigations in Biometry and Vital Statistics

(g) Special Investigations
Research on tuberculosis, hospital problems, etc.

Courses under the direction of Professor Raymond Pearl, assisted by a staff of seven associates.

University of Michigan. Graduate School

Vital Statistics (one hour, second semester); elementary methods in statistics; mathematical theory of statistics (a—introductory, b—advanced). Professor James W. Glover, and others.

Ohio State University. Graduate School

Demography. Open to students majoring in public health. Covers vital, social and sanitary statistics.

Social statistics. Application of statistical methods to social research. Miss Mark.

University of Pennsylvania. School of Hygiene and Public Health.

Social and vital statistics. Statistical principles connected with collection, presentation and interpretation of numerical facts, graphic methods, statistical control of public health work. Professor Robert Riegel.

Yale University. Department of Public Health.

Principles of vital statistics. Descriptive and analytic aspects of vital statistics. Life tables. Graphics. Assigned statistical problems in public health. Dr. L. I. Dublin.

New Employment Statistics. Arrangements have been made for securing monthly reports from a large list of manufacturing establishments in Massachusetts, showing numbers employed and total earnings, together with certain incidental information bearing on these two questions. The first canvass was undertaken during the month of September, 1922, and the results were highly satisfactory. In this work the State Department will coöperate with the United States Bureau of Labor Statistics and thereby avoid duplication in the collection of reports of this character. The information will be obtained in the first instance by this department, and copies of the reports will be furnished to the Federal Bureau for its use, in connection with similar data for other states, in compiling a report for the entire country. These monthly reports when compiled should serve as an important industrial barometer, showing the trend of employment and earnings of employees in the manufacturing industries.

Harvard Business Review. The Graduate School of Business Administration of Harvard University announces the publication, beginning in October, of the *Harvard Business Review*, a quarterly devoted to the theory and practice of business. The first number includes articles by Dean W. B. Donham, O. M. W. Sprague, Arthur S. Dewing, and other members of the faculty of the Graduate School, with contributions by business men such as George O. May, Frederic H. Curtis, and Julius Klein. The leading articles of the *Review* will be edited by an editorial committee of the faculty; the departments of business cases and book reviews will be in charge of a student editorial board.

Publications of the International Congresses of Eugenics. The First International Congress of Eugenics was held in London in 1912. The papers presented before this Congress are reported in two volumes entitled "Problems in Eugenics." There is also a supplementary pamphlet entitled "Catalog of Exhibits." American libraries can secure these reports directly from the Secretary-Treasurer of the Eugenics Research Association, Cold Spring Harbor, Long Island, N. Y.

The papers and reports of the Second International Congress of Eugenics, which was held at the American Museum of Natural History in New York in 1921, are now being printed by the Williams & Wilkins Company, Mount Royal & Guilford Avenues, Baltimore, Md. Volume I will be entitled "Eugenics, Genetics and The Family"; Volume II, "Eugenics in Race and State." Orders may be placed directly with the Williams & Wilkins Company for all publications of the Second Congress.

Besides these two volumes of scientific papers, there will be issued a supplementary pamphlet which will contain an account of the Congress, together with a Catalog of the Exhibits. Plans are in preparation for rebinding, under one cover, all accounts and descriptions of the exhibits.

The Pollak Foundation for Economic Research announces the publication of *The Making of Index Numbers: A Study of their Varieties, Tests, and Reliability*, by Professor Irving Fisher. This volume is a development of studies which were advocated and outlined by Professor Fisher at the meeting of the Statistical Association at Atlantic City, in December, 1920. Some of the preliminary findings were presented in the March, 1921, number of this JOURNAL. A special price will be made to members of the Association who send in advance orders.

The publication of Tables of $\sqrt{1-r^2}$ and $1-r^2$, by John Rice Miner, for use in partial correlation and in trigonometry, is announced by the Johns Hopkins Press.

Tables for Applied Mathematics to Finance, Insurance, and Statistics, edited by Professor James W. Glover, are now ready for delivery by George Wahr, Publisher (Ann Arbor, Mich.). These tables contain compound interest functions and logarithms of compound interest functions; life insurance and disability insurance functions; probability and statistical functions; and seven place logarithms of numbers from 1 to 100,000.

Annual Meeting of the American Statistical Association. The opening session of the Annual Meeting, to be held at Chicago, December 27–29, will be a joint meeting with the American Economic Association, on Wednesday afternoon, December 27. The general topic will be "The Outlook for 1923." Experts will discuss the level of prices and the cost of living, the trends in interest rates and investment, and wages and regularity of employment. Both the Economic Association and the Sociological Society will join with us in an evening session to hear the presidential addresses of the three associations. On Thursday, December 28, our Association will hold three sessions. The morning topic will be "Population Facts and the Malthusian Law." In the afternoon a round table will discuss "Standards and Methods of Statistical Instruction and Training." At this session the proposed plan, described in more detail elsewhere in this issue, for a series of standard reference books on statistical methods and practice will be presented for consideration by the Association. In the evening the Hon. Herbert C. Hoover, Secretary of Commerce, will deliver an address. Mr. Hoover presides over the most important statistical activities of the federal Government and is now carrying forward an active program for closer coöperation between the government and private business. The Association is fortunate in securing his presence at this most representative gathering of statisticians and economists. On Friday morning, December 29, a session will be devoted to "Employment Statistics and Allied Topics." The Association has had an active committee at work on this problem during the year, in coöperation with other agencies concerned, and important contributions by experts may be expected.

ROBERT E. CHADDOCK,
Secretary

Dr. Jacques Bertillon, who died at Paris on July 7 last, was one of the most eminent of the French demographers and was known as the father of International Nosologic Nomenclature. In 1893 he presented at the meeting of the International Statistical Institute, held at Chicago, a draft of a system devised by him and intended for international use as a standard for the statistical classification of causes of sickness and death. Such was the inception of what was known for awhile as the Bertillon Classification, but which by reason of its adoption by most of the countries of the civilized world has become the International List of Causes of Sickness and Death.

Prior to the adoption of the present system there was absolutely no international standard of nosological practice. Causes of death were classified by the several countries, and even by important political units within most countries, according to different systems. Obviously, the results were not comparable with each other. Dr. Bertillon's system of classification provided for the grouping under one title-heading of all causes of death reported in the same or in synonymous terms. He stipulated that by periodical revision his system was to be kept in conformity with the advancement of medical science. In consequence, there were held at Paris in 1900, 1909, and 1920 meetings of the Decennial Conference for the Revision of the International Nosologic Nomenclature. With each of these Conferences the number of countries participating has increased,

and as a result we have now in effect for the third decade a system which has given to the world comparable international mortality statistics.

It is of the utmost importance to the advancement of sanitary science that statistics of causes of death published by different offices be comparable with each other. There can be no doubt that Dr. Bertillon hastened by many years the uniformity of practice which resulted in the compilation of such statistics.

Dr. Jacques Bertillon was born in 1851 and was one of a distinguished family. He was the grandson of Dr. Achille Guillard, who was eminent both as a botanist and as a statistician. He was the son of Louis Alphonse Bertillon, a noted French statistician, and a brother of Alphonse Bertillon who invented the Bertillon System of Body Measurements for the use of criminal investigation bureaus. Dr. Bertillon served for twenty years as Secretary-General of the Commission for the Revision of the International Nosologic Nomenclature. For many years he was Director of the Paris Bureau of Statistics; he was a Chevalier of the Legion of Honor and an author of note. His works relate chiefly to the depopulation of France, the relation of marriage to longevity, and the effects of alcoholism on health and morals.

Dr. E. H. Downey, workmen's compensation insurance expert, was drowned at Avalon, N. J., July 9, 1922. During the past seven years Dr. Downey had been connected with the Department of Insurance of Pennsylvania, and he was widely known as a deep and thorough student of compensation insurance.

Dr. Downey was born in Iowa in 1879. He was graduated from Iowa State University, and later received his doctor's degree in economics from the University of Wisconsin. In connection with his university researches and his work with the Wisconsin Industrial Commission and the Insurance Department of Pennsylvania he published some very valuable monographs and scientific articles on accident compensation. Members of the American social science organizations will remember his very able address on workmen's compensation at the Annual Meetings in Pittsburgh last year. But Dr. Downey's greatest contribution was his untiring and intelligent activity in workmen's compensation administration. His untimely death is a very serious loss in the scientific development of legislation and administration in a field which yearly becomes increasingly complex.

Mr. Walter W. Stewart, Professor of Economics at Amherst College, has been appointed Director of the Division of Research and Analysis of the Federal Reserve Board. He takes the place of Dr. H. Parker Willis who recently resigned.

During the summer, prior to his acceptance of the position with the Federal Reserve Board, Dr. Stewart was engaged under the auspices of the Committee of the Plan of New York and Its Environs in the preparation of a preliminary plan and outline of economic portions of the investigations to be made in connection with the development of the proposed plan for Metropolitan New York.

Dr. W. A. Berridge, after two years as Instructor and Tutor in Economics at Harvard, has been appointed Assistant Professor of Economics at Brown. Dr.

Berridge has recently returned from England, where he has been collecting unpublished materials on the business cycle and on unemployment.

Dr. Rainard B. Robbins has returned to the University of Michigan as Assistant Professor of Mathematics after an absence of two years, during which time he was employed as Assistant Actuary in the New York Department of Insurance at Albany.

Mr. Arthur E. Albrecht, Instructor in Economics at the College of the City of New York, has been appointed to the position of economist in the New York City office of the State Department of Farms and Markets.

Mr. J. E. Vesely, formerly with the Wieboldt Company of Chicago, is now Statistician for the United States Gypsum Company.

Mr. Charles H. Wickens has been appointed Commonwealth Statistician of Australia to succeed G. H. Knibbs, C.M.G., who has become Director of the Commonwealth Institute of Science and Industry. Mr. Wickens has been one of the senior officers of the Bureau of Census and Statistics since its organization in 1906.

Professor Umberto Ricci has resigned as Chief of the Bureau of Statistics of the International Institute of Agriculture. The Institute's *Year Book of Agricultural Statistics* has been issued under the direction of Professor Ricci.

MEMBERS ELECTED SINCE AUGUST

Boudo, Carolyn, State Health Department, Albany, N. Y.
Cherington, Paul T., J. Walter Thompson Co., 244 Madison Ave., New York City
Fach, Herbert T., Rockefeller Foundation, 61 Broadway, New York City
Goldstein, Maurice L., Statistical Bureau, Metropolitan Life Ins. Co., New York City
Hoover, Herbert C., Secretary of Commerce, Washington, D. C.
Klein, Henry J., Statistical Bureau, Metropolitan Life Ins. Co., New York City
Lincoln, Edmond E., Western Electric Co., 195 Broadway, New York City
Lister, W. E., Actuary, Royal Indemnity Co., New York City
Richardson, I. G., Atlas Portland Cement Co., New York City
Roorbach, George B., Harvard University, Cambridge, Mass.
Schoenfeld, Wm. A., Marketing Research, Bureau of Agricultural Economics, Dept. of Agriculture, Washington, D. C.
Seligman, Mrs. Eustace, 161 E. 79th St., New York City
Smith, James G., Dept. of Economics, Princeton University, Princeton, N. J.
Stark, M. P., Washburn Crosby Co., Chamber of Commerce, Minneapolis, Minn.
Sturgis, Arthur, Consulting Economist, 712 Southern Bldg., Washington, D. C.
Thomas, Dorothy, London School of Economics, London, England
Thompson, Warren S., Scripps Foundation, Miami University, Oxford, Ohio
Wachsman, Charles, Statistical Bureau, Metropolitan Life Ins. Co., New York City
Wembridge, Harry A., Joseph & Feiss Co., Cleveland, Ohio
Wiesner, Dorothy, National Tuberculosis Association, New York City
Worsley, W. J., Statistical Dept., Edison Lamp Works, Harrison, N. J.

REVIEWS

Increase of Population in the United States 1910–1920, by William S. Rossiter. Bureau of the Census. 1922. 259 pp.

This publication, the first of a projected series of monographs upon various aspects of the census of 1920, is an attempt "to present a statistical picture of national progress," especially as mirrored in the changes of population and its various classes between 1910 and 1920. The statement of purpose suggests that the monograph might be compared with Tucker's *Progress of the United States in Population and Wealth in Fifty Years as Exhibited by the Decennial Census* (1843), a very able attempt by a private scholar to elucidate the results of the first six censuses which the government had printed with no effort at interpretation. It might also be compared with various chapters or sections in later census reports after some influence, possibly that of Professor Tucker, had led the officials, beginning with DeBow in 1850, to present the main results in words as well as figures. Among the publications of the latter class these may be mentioned:

In the Tenth Census: Population, Volume I—Walker and Gannett's "General Discussion of the Movements of Population 1790 to 1880," especially pages xi–xli.

In the Eleventh Census: Population, Part I—Porter, Gannett and Hunt's "Progress of the Nation 1790 to 1890."

In the Twelfth Census: Population, Volume I—Hunt's discussion of the "Statistics of Population," and in the Supplementary Analysis, various chapters by the writer and others.

In the Thirteenth Census: Several chapters introducing the tables dealing with the different topics.

One might go farther afield and compare this monograph and its successors with the introductory volumes in several foreign censuses. A careful study of the degree and way in which the census figures of population have been interpreted in different countries and at different dates by the office responsible for them would probably be worth while. But such a study would exceed the scope of a review and call for more space than is now available. Suffice it to say that the present monograph has been prepared in a manner somewhat different from its forerunners. Heretofore such writing has been done either by an outsider making no use of office resources (Tucker) or by a member of the staff. The writer of this analysis is one who won high position and reputation within the office and then withdrew to more inviting fields, but who did not end his intimate association with his former colleagues of the Census Bureau.

His work is a product of coöperation between an outside expert and the office which finds its only prototype in the manufactures census of 1900. It is the first fruit of a policy introduced in the census of 1920 at the suggestion of the Advisory

Committee of the American Statistical Association and the American Economic Association. It appeared that the mandate of Congress for the completion and publication of the preceding census within three years, a mandate without a sanction, had not been obeyed. Therefore the advisory committee urged that rapid production was after accuracy the most important duty and opportunity of the office, that to assure this result there should be no delay for the preparation of text other than what was needed to make the meaning of the figures clear, and that the Bureau should arrange for the publication later of special studies or monographs to aid in the interpretation of its figures. The first of these prepared by the chairman of the joint Advisory Committee is now given to the public.

With these explanations the following summary of some of its main points, with running comments or criticisms in the few cases where I differ, will stand in its proper perspective.

The population of the United States in 1920 exceeded that in 1910 by 13.74 million. To determine the number of new members of the population at the later date the monograph adds to this 13.74 million 11.24 million to balance the number of persons enumerated in 1910 who died between that year and 1920, and 2.28 million to balance the number who left the country in the same decade and were alive at its close.

To estimate the number who died, 1910–20, among those enumerated in the earlier year the office has applied the annual death rate of the registration area in 1910 to the population of the country in that year, the rate for the population above one year of age in the registration area in 1911 to the survivors throughout the country in that year, and so on. One of the assumptions underlying this method, namely, that the death rates in the registration area give the closest approximation available to the death rates in the country, is quite far from the truth. For the proportions of the sexes and ages in the population of the registration area may be different from those in the population of the United States, and the proportions of the two main races certainly are different. Thus the present method assumes a death rate of 14.958 per 1,000 in the population of the United States in 1910 because that was the rate in the registration area. But the death rate of whites in the registration area (14.573 per 1,000) applied to the whites in the country yields 1,191,080 deaths, and the death rate of colored in the registration area (24.202 per 1,000) applied to the colored in the country yields 247,836 deaths, a total of 1,438,916 estimated deaths in the country, indicating a rate of 15.645 instead of 14.958 per 1,000 for the entire country in 1910. In other words, after a correction for race it appears that the death rate proper to use exceeds that employed by 0.687 per 1,000 or 4.59 per cent. With this correction the estimated number of deaths in the decade is raised by 516,000, and the population in 1920, estimated by adding to the survivors of those enumerated in 1910 all natives under 9 and 17/24 years (the interval between the censuses of 1910 and 1920), and the survivors of all immigrants becomes 105,400,000 instead of 105,920,000. The result is of importance because the figures in the monograph point toward omissions of 210,000 in the census of 1920, while my substituted figures suggest that the enumeration of 1920 found 310,000 more people than would have been expected.

The rate of decennial increase in the population of the United States has been falling since 1860, but the regularity of this change is obscured by omissions undoubted and serious at the census of 1870 and probable though smaller in amount at that of 1890. The rates of decennial increase since 1900 after correcting for the fact that the enumerations were not quite ten years apart were 21.3 per cent between 1900 and 1910, and 15.4 per cent between 1910 and 1920, so that the rate in the later decade was less than three-fourths of that in the earlier.

This checking of the growth of population affected the great divisions of the country very differently. If we neglect for a moment the far West, that is, the mountain and Pacific divisions, and consider only the North and South, it appears that between 1900 and 1910 the South increased about one-seventh faster than the North, but between 1910 and 1920 the North increased about one-fifth faster than the South. In the earlier of these two decades the states west of the Mississippi increased seven-tenths faster than those east of it, but in the later they increased only one-seventh faster. To present the changes in a somewhat different way, the states of the mountain and Pacific divisions increased in the second decade of the country less than half as fast as in the first, the southern states increased about two-thirds as fast, and the northern states increased four-fifths as fast. This closer approach to the earlier rate of growth in the northern states with a high proportion of urban population suggests, what other figures establish, that the cities have retained their former rates of growth more nearly than the rural districts have. In fact, the increase of American cities in 1910–20 was at a rate nearly three-fourths of that in the preceding decade; the increase in the rural districts fell to less than one-half of the rate ten years earlier.

May I here repeat a suggestion of many years ago that the census definition of an urban district or city, namely, an incorporated place containing at least 2,500 inhabitants, is provisional and temporary. The main distinction between city and country, I take it, is that in the country more food and in the city less food is produced than is needed to support the people who live within the area. If this be so, the line of separation should be based not on total population but on density. For example, there is no good ground in my opinion, except the fact that we lack a practicable and better definition, for calling Ithaca in Tompkins County, New York, a city and Groton in the same county rural because the former has 17,000 people and the latter only 2,235. For the average density of population in Groton is 3,788 and in Ithaca 3,721 persons to a square mile. Perhaps in future decades, when the area and population of all small divisions of the United States have been measured, we shall define rural districts as those with so low a density of population that they produce more food than they consume, and cities as those districts with so high a density of population that they consume more food than they produce. An intermediate class, composed of districts where both agriculture and other occupations are important sources of a livelihood, may be necessary.

In comparing the growth of city and country one may define the urban population of 1910 or 1920 as the population of all incorporated places which then had at least 2,500 inhabitants, ignoring the fact that several hundred places had less than that population in 1910 but more in 1920. Or one may treat as urban at

both censuses those places and those only which in 1920 (or 1910) had at least 2,500 inhabitants. Under the former method American cities grew, 1910–20, nine times as fast as country districts; under the latter and more accurate method they grew less than five times as fast. This second method, however, is not above criticism, for it rests on the assumption that the boundaries of all cities were the same in 1920 as in 1910, although one may feel sure that a significant number annexed territory during the decade and thus swelled their real increase. Under the method of a diminishing rural area the gain in the rural districts, 1900–10, was more than one-fourth, and 1910–20, little more than one-ninth of the increase in the country as a whole. If the large group of northern states from Maine to North Dakota and south to Mason and Dixon's line, the Ohio, and the southern boundary of Missouri and Kansas be regarded as a unit, the number of persons living outside of incorporated places with at least 2,500 inhabitants was about 175,000 less in 1920 than in 1910.

The excess of males which had been noted in the United States for a century rose slowly from 1820 to 1860 when it was nearly three quarters of a million, fell in the following decade by 300,000 as a result of the Civil War and the concurrent check to immigration, then rose for forty years until in 1910 it was about 2,700,000. But in the last decade of war and slackened immigration it fell again by 600,000. In 1920 the excess of males was still above 2,000,000 and greater in amount than at any date before 1910. But the ratio of excess was less than in 1890 or 1900.

About age little is said directly in the monograph, and occasionally the writer seems not to allow enough for changes in age composition. Thus, he tries to explain the larger proportion of widows in 1920 as compared with New Hampshire figures of 1773 by supposing that before the Revolution "the marital relationship was held to be more desirable" and that it was then more difficult for widows to support themselves (page 147). Although these were probably factors I suspect that an influence at least equally powerful may be found in the fact that the average age of the population in 1920 was fully ten years greater than in 1773, and that widowhood being a function of age naturally increased with it.

Mr. Rossiter suggests that the extinction of the North American Indian (sc. in the United States) "has been averted by increasing intermarriage" (page 137). Does not this depend upon the definition of extinction? Few would offer evidence to show that the amount of Indian blood in our population has diminished. But it may be true, nevertheless, that the number of persons with obvious Indian characteristics is decreasing, and intermarriage would not avert but rather contribute to that kind of extinction. Or if one means by an Indian a person living in tribal relations under Indian law and custom, the evidence for approaching extinction is even more convincing.

Notwithstanding these few points at which I am disposed to take issue with the writer's conclusions, or rather with the office methods or practice upon which he stands, the monograph as a whole commends itself as careful, thorough, discriminating, and above all, readable and significant. It will arouse a new and wider interest in the meaning of census figures and of the changes in population which they mirror and reveal.

We may admit for it and hope for its successors that the policy exemplified by the series, as the Director claims in his prefatory note, "is clearly a long step in advance in the effort to make the decennial census of as much practical value to the Nation as possible" (page 7) and yet ask whether this effort goes more than half way toward the ideal. The Bureau of the Census ought to be the best place in the United States for the training of statisticians. A forceful and competent director, interested in securing that result, who had a reasonably permanent tenure and who was able to get backing from Congress, might make it so. In the twenty years of its life as a permanent office the Bureau has had directors of various types, but none, I believe, who has made this a real part of his job. Until his work is more nearly divorced from politics or until the assistant director becomes the permanent and recognized head of the scientific personnel and the director's chief of staff in all matters scientific or professional, the organization of the Bureau of the Census on the scientific side will continue to be seriously defective. Work like that I tried to do in 1900 and like Mr. Rossiter and his followers are doing in this series may salve the defect for an emergency; they cannot really cure it.

WALTER F. WILLCOX

Cornell University

Wages and Hours in Anthracite Mining. National Industrial Conference Board, Research Report Number 47. New York: The Century Co. 1922. 67 pp.

This report is compiled from questionnaires submitted to 47 per cent of the operating anthracite mining companies. It covers about 60 per cent of the employees in the anthracite mining industry. Women workers, boys, foremen, salaried and clerical workers are excluded from the report. Moreover, all employees who did not work within two days of the full time offered by the employers in the semi-monthly periods studied were excluded from the report. This was done because the mining companies contend that to include all wage-earners irrespective of the time worked "would give a distorted picture," since miners frequently shift from mine to mine or from job to job within the same mine. This turnover increases the total number on the payroll in any given period and lowers the apparent average earnings. The report itself recognizes the weakness of this procedure, for it says:

> Obviously, the comprehensive way of dealing with the wages and hours of work situation in the anthracite mining industry would be to secure and analyze pertinent data for all mine workers, properly classified according to the number of days and hours worked in each pay period, and to give each group its proper weight in the final determination of the average conditions in the industry in respect to wage earnings and working time. This would also permit of definite evaluation in comprehensive and scientific manner of such peculiarities of labor turnover and payroll records as are claimed to exist in the industry. . . . In presenting the results of its investigation the Board therefore reserves judgment as to their adequacy in representing general wage-earning conditions in the industry until more evidence is available regarding the effects of excluding the shifting or part-time workers.

Nevertheless it is argued (p. 9) that the study "affords a fair picture of wage-earning conditions in the industry."

The schedules compiled were for the last half of each of the following months: June, 1914; October, 1920; March, 1921; June, 1921; and October, 1921. Taking June, 1914, as the base, average hourly earnings of all wage-earners covered except contract miners had risen to the index number of 266 by October, 1921; for contract miners, to 258. Semi-monthly earnings for all workers (contract miners excluded) had risen to 258 by October, 1921; for contract miners, to 243. The average weekly hours worked by all wage-earners (contract miners excluded) decreased from 51.6 hours in June, 1914, to 48.2 hours in October, 1921; for contract miners, from 42.4 hours to 39.9 hours.

These figures challenge one's attention. The award of the Anthracite Coal Commission in August, 1920, granted an increase in the gross tonnage rate for contract miners of 65 per cent over the 1916 rate. Since the 1916 rate was 7 per cent over the rate obtaining in June, 1914, the base period used by the Board, the tonnage rate for October, 1921, would be 76.5 per cent higher than the rate obtaining in June, 1914.[1] Yet the Board finds that the hourly earnings of contract miners had increased 158 per cent and the semi-monthly earnings 143 per cent. It is true that changes in the tonnage rates do not accurately represent changes in earnings because deductions for helpers and supplies have not changed in the same proportion.

The Board finds that the earnings of contract miners in June, 1914, averaged 41 cents an hour. The Board assumes that in 1914 contract miners worked on the average 8 hours per start or day (p. 12). This would make the daily earnings $3.28 for June, 1914. Compare this with the 1914 daily wage of $5.26 for contract miners as given in a table compiled from wage agreements and a memorandum submitted to the Conference of National Labor Adjusting Agencies of November, 1918.[2] On the other hand we find that the semi-monthly earnings of contract miners for October, 1921, as given by the Board were $91.37, while the Bureau of Labor Statistics finds that the semi-monthly earnings for January, 1922, were $86.60.[3] Considering the difference in periods as well as firms covered there appears here to be substantial agreement.

Turning now from the contract miners to the rest of the wage-earners we find a fairly close correspondence between the data presented by the Board and other available data. The following table compares the hourly rates compiled from the wage agreements and the memorandum submitted to the Conference of National Labor Adjusting Agencies with the hourly earnings given by the Board. The Board's figures are for June, 1914, while the memorandum figures are for 1914.

	N. I. C. Board	Memorandum
Skilled labor (outside)	.265	.277–.303
Common labor (outside)	.183	.169–.203
Common labor (inside)	.220	.225–.251
Skilled labor (inside)	.275	.277–.303 (company miners)

[1] *Bulletin of the U. S. Bureau of Labor Statistics*, No. 279, p. 12.

[2] *Wages in Various Industries and Occupations*, Bureau of Applied Economics, Bulletin No. 8, p. 6. Mr. Ash's estimate, however, made in the *Labor Review* for October, 1920 (p. 104), corresponds much more closely with the Board's figures.

[3] *Monthly Labor Review*, May, 1922.

For the post-war period we find that the Board's figure for hourly earnings of skilled inside labor for October, 1921, is 0.677, while the data furnished by the Bureau of Labor Statistics for January, 1922, give 0.686 as the hourly earnings.[1] For common inside labor the Board's figure for October is 0.596, while the Bureau's data for January, 1922, give the hourly earnings as 0.599.[2] The semi-monthly earnings of skilled labor as shown by the Board were 9.5 per cent higher than the semi-monthly earnings shown by the Bureau. This is due to the fact that the average hours actually worked were considerably higher as shown by the Board than as shown by the Bureau. In the case of common labor the hours worked were about the same in the two studies, and therefore since the hourly earnings were similar the semi-monthly earnings were likewise similar.

The wage increases of contract miners especially (contract miners constitute about 30 per cent of all anthracite wage-earners) as shown in this report appear to be somewhat greater than other available data would indicate. This may be partly due to the questionnaire method which has been criticized by Mr. Douglas in previous issues of this JOURNAL. In part it may be due to the fact that the base period used by the Board (June, 1914) was far from normal, as the present reviewer has elsewhere attempted to show.[3] Although a study of wage rates would not be affected much by using such a base, a study of earnings using an abnormal period for a base would present a distorted picture. Again, it does not appear quite fair to compare the wages from October, 1920, to October, 1921, with those for June, 1914, without showing the earnings in the intervening period. The great increase in anthracite wages came in 1920. The earnings of anthracite labor in 1920 and 1921 were far more favorable than those of earlier years, particularly the years prior to 1918.

ALVIN H. HANSEN

University of Minnesota

Report on Productive Industries, Railways, Taxes and Assessments, Waterways, and Miscellaneous Statistics of the Commonwealth of Pennsylvania for the Year 1920. Department of Internal Affairs. Compiled by the Bureau of Statistics and Information, M. Hoke Gottschall, Director, Harrisburg, Pa.: J. L. L. Kuhn, Printer to the Commonwealth. 1921. 1,040 pp.

In a "foreword" to this report the Secretary of Internal Affairs, James F. Woodward, makes the following statement in the opening paragraph:

> In the year 1872 the first official report on industrial statistics for the Commonwealth of Pennsylvania was issued by the "Commissioner of Statistics" and the presentation of this report covering the year 1920, therefore, becomes the 48th Annual Report of the Bureau of Statistics and Information. The records of the Department of Internal Affairs show a complete set of these Annual Reports and they portray a marvelous development of industry in Pennsylvania.

[1] *Monthly Labor Review*, May, 1922. The above figure for skilled labor was obtained by taking a weighted average of the hourly earnings of blacksmiths, bratticemen, machinists, masons, company miners, timbermen, and trackmen as given by the Bureau.

[2] For common labor a weighted average of the hourly earnings of cagers, car-runners, drivers, and laborers was calculated.

[3] "The Buying Power of Labor During the War," this JOURNAL, March, 1922.

As indicated by its title, the report is of wide scope, relating, as it does, not only to the industrial activities of the Commonwealth, but also including "miscellaneous statistics reflecting activities of the various departments of the Commonwealth for the year 1920." Virtually, therefore, it fulfils a two-fold purpose—that of recording the progress of the industries of the Commonwealth, and that of serving as a statistical abstract of Pennsylvania, analagous, in general plan, to the Statistical Abstract of the United States. Thus in a single volume are brought together the principal statistical data relative to the Commonwealth, rendering it unnecessary for one to consult numerous official reports in order to gain a general knowledge of the resources and development of the state. In connection with each statistical table, not obtained from *original* sources by the Department of Internal Affairs, due acknowledgment is made of the "secondary source" from which the data were obtained.

Aside from the brief foreword (3 pages) and a few paragraphs accompanying several of the miscellaneous tables, the entire report consists of statistical tables. It is to be regretted that the report did not include a description of the methods employed by the Bureau of Statistics and Information in securing data from original sources, and at least a few pages of interpretative text, with a discussion of the more significant results of its inquiries. Prior to the issuance of the report, summaries, in text form, of the principal facts with reference to the industrial activities of the Commonwealth were issued as "press releases," which, if they had been incorporated as an introduction to the report, would have added materially to its interest, at least to many persons who have little, if any, interest in a mere collection of statistical tables. In the foreword, it is true, allusion is made to "the more than 5,000,000 separate tabulations herewith presented," but the reader is left in disquieting doubt as to what is meant by a "tabulation," as the number of separate tables can be expressed with not more than two digits, and even of separate totals a rough estimate does not reach to a half-million.

Because of the fact that the report consists almost wholly of tabular matter, a critical review must confine itself primarily to the form and arrangement of the tables. These conform to accepted practice in official publications, except that the titles are not sufficiently explicit to indicate adequately the nature of the data presented therein. This criticism is particularly true of the tables of "industrial statistics" constituting Parts I, II, and III of the report and comprising over 600 pages. Only by an examination of the box headings can one determine that the data in these tables relate to "industries," and even then only after reading through several pages of side headings does one become aware that the data do not relate wholly to manufacturing industries but include also building and contracting, mines and quarries, and public service. Part III, entitled "Detail Industrial Tabulation for 1920 by Counties and Selected Municipalities," consists of two tables of similar form, together occupying 545 pages. As no continued titles or page headings are given, the first page of the second table, which has reference to municipalities, cannot be readily located. The table of contents might be of service in assisting one to locate the various tables in the report were it more detailed, but it is merely a list of the titles of the nine "Parts" of the report and does not even include references to the initial pages of the respective

parts. A complete list of the tables with page references would enable the reader to turn to any table without difficulty. Usually in voluminous governmental reports consisting principally of tabular matter (as, for example, the United States census reports) a table of contents and a list of tables, both with initial page references, appear immediately or closely following the title page.

There is an index to the report of which no mention is made in the foreword or table of contents, so-called. For one not already familiar with the subject-matter of the report the index does not serve adequately as a reference guide to particular tables. For example, there is among the large number of miscellaneous tables one entitled "Number of arrivals and departures of vessels," but there does not appear to be any reference whatever in the index to "vessels," "arrivals," or "departures," nor any means of locating this table other than by laboriously turning the pages until page 898 is reached.

No opinion as to the accuracy of the data presented in the report would be justified unless the critic had first inquired into the methods employed in the collection of information from the original sources, and in editing and tabulating the returns; but a close examination of certain of the tables leads to some question as to whether or not requisite care was observed in editing the report. By way of illustration: In the first table in the report, which bears the title "State Summary," industrial statistics for the several years 1916, 1917, 1918, 1919, and 1920, are presented in parallel columns and, presumably, these data are strictly comparable; yet it appears that the "number of plants operated" in 1920 was 26,077, while in 1919 there were only 20,888, and that there was an increase, in a single year, of 5,189 plants operated, or nearly 25 per cent, which increase any one reasonably familiar with industrial conditions in Pennsylvania during these two years would immediately question. The explanation of this inordinate increase in the number of plants operated is found in the following statement which appeared in a press release issued by the Department on October 21, 1921.

> According to the preliminary figures announced by Secretary of Internal Affairs Woodward, there were 26,076 industrial establishments reporting to the Pennsylvania Department of Internal Affairs in 1920. In the preceding year there were 20,888, and included in the increase last year were numerous small plants which had never before been called upon to make reports and which affected the general production valuation total but little. Of course newly established plants were also included in the increase.

Evidently the press release was issued several months prior to the date on which the report was issued in print, as a "letter of transmittal" appearing in the report was dated December 1, 1921. There was, therefore, ample opportunity to include in the report (at least as a note to the table above referred to) a statement explanatory of the increase in the number of plants operated, to which plants the industrial statistics, comprising considerably more than one half of the report, relate.

Objection might be raised to the publication of data relative to wages paid, capital invested, and total value of products in round hundreds of dollars when the amounts are comparatively small. Numerous items and totals of less than $1,000 appear in the tables of industrial statistics. Readers of the report are entitled, at least, to know whether the tens and units were eliminated when the

questionnaires from individual reporting agencies were edited or when the tables were assembled for printing, but the report furnishes no explanation as to the method employed.

In general it may be said that the report under review should prove to be a very useful contribution to the rapidly growing list of official statistical reports periodically issued. Other states which do not issue a compendium of information relative to their industries and resources might well follow the example of Pennsylvania in this respect. Hardly any other "medium for disseminating information the public should have" could prove more effective than a volume of this character.

ROSWELL F. PHELPS

Wealth and Taxable Capacity, by Sir Josiah Stamp. London: P. S. King & Son, Ltd. 1922. 195 pp.

Sir Josiah Stamp's new book contains the Newmarch Lectures for 1920–21, and represents an "attempt to outline some of the elementary principles involved in the valuation of national capital and income, and the determination of their distribution and their relation to prices, taxation, and public debt." There are six chapters, dealing respectively with the National Capital, the National Income, the Distribution of Income and Capital, the Limits of Taxable Capacity, the Effect of Changing Price Levels upon Profits and Wages, and the Effect of Changing Price Levels on the Burden of Public Debt. The book is short and, in the words of the author, "suffers from the defects of the lecture form." It is difficult reading at points; the author is forced in many places to state conclusions rather dogmatically and to forego explanations which would be particularly welcome to foreign readers. But the essential questions involved are dealt with in a most helpful way, and in many respects the book gains from its brevity. One sees the forest better because some of the trees are omitted or sketched in the most impressionistic way, and the limitations under which the lecturer spoke only serve to emphasize his unusual mental vigor and statistical ingenuity.

It is really an essay in financial arithmetic, rather than a statistical discussion of the ordinary kind. Most of the figures are approximations subject to wide margins of possible error. The treatment is beset at every hand with technical and theoretical difficulties, which can be bridged only by the exercise of great ingenuity. Difficult questions of ethics and political psychology are involved. Yet these questions must in some way be answered. Our financial leaders must have quantitative ideas about the taxable capacity of their own and other countries, particularly that of Germany. British leaders, at least, must reach some conclusion about the approximate yield of a capital levy, or a levy upon war-made wealth. We must somehow form a conclusion about the distribution of the national dividend and decide whether the trend of change is encouraging or discouraging, wholesome or the reverse. Taxpayers and taxgatherers are particularly concerned at the present time with the effect of changing price levels upon profits and the public debt; and the last two lectures are unusually valuable, not because they represent a rounded and final treatment of this subject, but because

they break a path through a statistical wilderness which somehow must be explored and mapped. It is in questions of this kind that Sir Josiah Stamp's versatility and high-powered mind are most helpful.

The book is full of useful cautions pointing out the pitfalls that must be avoided in dealing with national wealth, national income, and taxable capacity. It is full also of important conclusions, such as the finding "that the ordinary person of today is four times as well off in real commodities as the person in the corresponding stage in the scale in the beginning of the nineteenth century"; or that if all the incomes over £250 per annum were pooled and divided, it would add an amount of approximately five shillings a week, or perhaps a little more, to each British family. The conclusions relating to the effect of change in price levels upon profits are unusually suggestive. The author concludes that for short periods the fluctuation in profits may be roughly taken at two-thirds to three-fourths of the corresponding fluctuations in turnover, as represented by banking and foreign trade statistics, which reflect both quantities and prices. In coal mining a change in price represented by one point was found to be accompanied, on the average, by a change in profits of approximately three points. In the case of merchants' profits, however, "the fluctuations in total profits were almost entirely those due to the amount of sales and the results were unaffected by price changes."

These statistical results are supplemented by a theoretical examination of the relation between price changes and profits, from which the following equation is deduced: $\text{Profit} = \frac{1}{R} T P_2 + S(P_2 - P_1)$, in which R is the rate of profit, T the turnover, P_1 and P_2 the two levels of prices, and S the stock. The inference from this equation and the general analysis on which it is based would seem to be that total profits or losses in any year are equal to the change in price multiplied by the average stock carried, plus the normal profit on sales. This conclusion corresponds with American experience that in times of rapid price changes one of the most important elements of profits is the increase or decrease in the value of stocks on hand. But it is in apparent conflict with the statistical finding regarding merchants' profits, where the fluctuations were found to vary almost entirely with the amount of sales. The author at this point, however, is particularly occupied in driving home the useful truth that because of the vital connection between stocks, price changes, and profits, it is essential that deflation shall take place gradually, so that shrinkage in the value of stocks may represent merely "a reduction and not a total cancellation of the ordinary profits."

Statistical studies of the kind discussed in these lectures are not only helpful, they are absolutely necessary to the formulation of any sound financial or political program. But when one considers the estimates which must be made, and the psychological or political factors which must be allowed for although they cannot be measured statistically, the conclusion is borne in upon the reader that they are more practical than scientific. For instance, the computation of national wealth is arrived at by capitalizing statistics of income. The latter themselves are faulty and must be corrected in highly important ways. When this has been done the appropriate rates of capitalization must be applied. Anyone

who is familiar with the essential uncertainty attaching to the proper rates of capitalization and the extent to which they vary, not only from industry to industry but for particular business units within each industry, must doubt the scientific significance of the aggregate. One may—indeed, one must—figure on changes in this aggregate for certain purposes; but the reviewer, at least, would be much better satisfied if we confined the statistics to measurements of change and called the base, or starting point, x.

The difficulties and, to a certain extent, the unreality of these huge aggregates are well illustrated by W. I. King's criticism [1] of Stamp's treatment of interest. Stamp suggests that interest upon war debt should be deducted from the national income (based upon taxable income) because unless we do this "the bigger the debt, the bigger the income." King replies that "the author apparently overlooks the fact that the borrower, by obtaining goods when he wishes them, may thereby gain an additional service equalling in value the interest that he pays. Is it not reasonable to include the value of this additional service in the total of the national income?" King classifies the victory over Germany, or the costs thereof, along with payments for groceries and theater tickets, as a consumptive expenditure. He not only does that but he suggests that the service for which we pay interest is a real service which should be included in the national income *when the interest payment is made*. If, as a people, we borrow for purposes of consumption today, the interest which our children must pay for our fun really increases the national income. The greater our extravagance in this generation and the deeper in debt we go because of it, the larger the money income of the next generation will be.

It seems absurd: but the absurdity is not Dr. King's or Sir Josiah Stamp's. Both are in essential respects right. Waiting is a service; interest is the payment for waiting; and our usual economic definition of income includes service. But the concept of income which includes this kind of service and excludes the thousands of services which happen not to be measured in money or the equivalent of money is essentially misleading. I should add that Dr. Stamp recognizes the logical flaws inherent in the usually accepted definition of national income.

T. S. Adams

Statistical Year Book of the Kingdom of Siam. 1921.

This is a most extraordinary statistical publication and reflects the profound influence of British administrative methods on alien forms of government in the direction of progress and civilization. The *Year Book* is published by the Department of General Statistics, Ministry of Finance, in an English edition dated December, 1921. Its main divisions consider: (1) meteorology; (2) population; (3) finance; (4) foreign trade and commerce; (5) posts, telegraphs and telephones; (6) state railways; (7) education; (8) justice; (9) agriculture; (10) opium; (11) mining; (12) prices; (13) cultivated areas under rice; and (14) weights and measures. In other words, the *Year Book* reveals the state of the Siamese nation in all its essentials as required for legislative and general purposes. The presentation

[1] *American Economic Review*, Sept., 1922, p. 532.

of the facts is made with admirable clearness and due consideration of scientific requirements.

The census figures are for the year 1911–12, but a calculation is included of the population for intercensal years from 1901 to 1921, the figure being now placed at 9,121,000. The density of population is 17 per square kilometer, varying from 5 for one district to 214 for another. The sex distribution for the year 1911–12 reveals 4,122,168 males and 4,144,240 females, or a total of 8,266,408. The census is by age and sex, but unfortunately for age divisions somewhat at variance with accepted methods.

The economic progress of the country is perhaps revealed most clearly in the governmental revenue and expenditures, the income having increased from 17,334,000 ticals in 1894–95 to 96,066,000 in 1919–20, while the expenditures have increased during the same period from 12,487,000 to 82,560,000 ticals. Among the items of revenue of special interest is the income from opium licenses, given as 2,443,000 ticals in 1894–95 and as 23,222,000 ticals in 1919–20. Among other taxes of special sociological interest is the income derived from the so-called gambling farm, given as from 2,283,000 ticals in 1894–95, reaching a maximum of nearly 7,000,000 ticals in 1904–05, and declining slowly until in 1919–20 the revenue was only 760 ticals, gambling having been abolished first in the provinces and later, as recently as 1918, in the city of Bangkok. Likewise the revenue from the lottery farm, which at the beginning of the period was 1,037,000 ticals, and almost 3,000,000 ticals in 1916–17, ceased entirely in 1919–20.

Encouraging is the increasing expenditures for local sanitary purposes, which in 1904–05 amounted to 1,338,000 ticals, and in 1919–20 to 2,271,000.

The outstanding total national debt of Siam increased from £1,000,000 in 1905–06 to £7,312,560 in 1920–21. The *Year Book* includes also interesting observations on the Siamese Government 4½ per cent Sterling loan of £1,000,000 issued in 1905, of a similar loan of £3,000,000 issued in 1907, and of a loan contracted with the Government of the Federated Malay States for the construction of the Southern Railway, in 1909, of £4,750,000. The finance section also includes an illuminating note on the Gold Standard Reserve Fund, with tables on the exchange operations, by single years, from 1911–12 to 1919–20. This is amplified by a table in detail of exchange quotations for demand drafts on London, Singapore, and Hongkong, by months, from 1902 to 1921; and further by extended tables of the price of bar silver in London from 1900 to 1920. This is followed by a highly interesting note on the currency of Siam, also amplified by extended tables on the number and value in ticals of subsidiary coins imported, held in the central treasury, and in circulation, on March 31, 1921, and covering the period March 31, 1880–81 to 1920–21.

Of special interest is a brief table referring to the progress of savings banks, which in 1916–17 had 1,663,365 ticals on deposit, and by 1920–21 had 2,780,760 ticals.

The working of the state railways is dealt with in full detail for every year since 1897–98, the gross earnings having increased from 232,000 ticals during the first year to over 9,651,000 during 1920–21. The tables in detail contain information regarding the number of passengers carried, the receipts for luggage, the receipts

from passenger and freight traffic, and gross earnings per kilometer of line and per 100 train kilometers. There are full details regarding particular items of freight, differentiating goods and livestock, as well as different classes of passengers. The total passenger traffic increased from 276,747 in 1897–98 to over 5,189,400 in 1920–21.

The tables on education are suggestive of the progress which is being made, the number of schools having increased from 391 in 1915–16 to 620 in 1919–20; and the number of scholars, from 29,218 to nearly 50,000 during the same period. These statistics are for government schools only; but aside therefrom in 1919–20 there were 2,355 local schools, with 120,454 additional pupils.

The judicial statistics are admirably condensed and sufficient for ordinary purposes, clearly differentiating the nature of the crimes, by local divisions of the country, for each of the last five years, and summarized for the five-year period. In the entire Kingdom of Siam, among more than nine million people, in the year 1919–20 there were 310 cases of wilful murder, but in addition thereto there were 194 cases of manslaughter. About one-fourth of the crimes were on account of theft, and about one-sixth on account of contravention of gaming laws. The judicial statistics include tables of the number of persons convicted in the First Instance, classified by the nature of the sentences pronounced; also statistics of the work of the Court of Justice of First Instance in civil cases and their disposition.

Under agriculture there are some interesting notes on the registration of mortgages, the number of deeds issued and filed having been 565,454 during the year 1920–21. The registered mortgages are classified according to the nature of the land, the area, the number of holdings, and the value in ticals. Of special interest in this connection are the areas under certain crops, differentiating rice, cotton, tobacco, peas, sesame, pepper, and maize. The most important crop, naturally, is rice, the area of which increased from 13,559,000 rai in 1916–17 to 15,293,135 rai in 1920–21. Cotton growing seems to be increasing, for in 1916–17 the area grown was 29,070 rai against 40,867 in 1920–21. There are additional tables of the number of cocoanut trees, in bearing and not in bearing; the number of elephants, which practically throughout the whole kingdom decrease from year to year, there being in all Siam at the end of 1921 only 6,818 elephants left. The number of buffaloes for the same period was given as 2,603,823.

The opium traffic in the Kingdom of Siam is still of very considerable extent. The opium law of 1920–21 prohibits the smoking of opium except by licensed smokers. Since 1908–09 the control of the sale of opium has been in the hands of the Royal Government. The total value of opium purchased by the Opium Department has increased from 2,116,174 ticals in 1907–08 to 4,064,086 ticals in 1920–21. There are at present 886 retail opium shops in Siam.

The mining statistics are limited to tin, the price of which has almost doubled during the last fifteen years, though declining slightly during the year 1920–21.

The statistics of prices are on the basis of index numbers to reflect the fluctuations in the prices of certain stable commodities. The basic prices are those prevailing in 1914–15, which conforms to modern usage. On this basis the price index number for foodstuffs during 1920–21 was 181; for beverages, 268; and for

both combined, 204. The price index number for imported cotton goods increased from 100 in 1914–15 to 276 in 1920–21.

Wages show practically a stationary condition, having undergone but a slight change for ten of the principal employments.

The concluding tables of the report have reference to areas cultivated in rice as graded for revenue purposes.

The book is beautifully printed, on good paper, is of convenient size, and is in every respect a most valuable addition to the growing official literature summarizing the results of governmental operations and inquiries. It must ever be a matter of regret that this admirable precedent set by foreign countries should never have been imitated in a similar manner by even a single American state. The progress of any government, in its final analysis, depends primarily upon a clear grasp of governmental facts, and this is impossible without a condensed statement of the vast bulk of statistical information which is now being gathered by practically every civilized country in the world.

The Ministry of Finance of Siam is setting an admirable example for other nations to follow. The publication of the reports in English makes accessible a wealth of information of which much would otherwise go entirely to waste. The *Year Book* is a useful companion volume to the annual report of the Financial Adviser on the Budget of the Kingdom of Siam, whose report for the year 1920-21 has also recently been issued in the English language.

F. L. HOFFMAN

United States Life Tables 1890, 1901, 1910, 1901–1910, prepared by James W. Glover. Published by the Bureau of the Census, Washington. 1921. 496 pp.

This volume, which is divided into eight parts, constitutes a complete report on the mortality tables constructed in connection with the census of 1910. The introduction gives a general description of the entire contents of the report and outlines the purposes in view in presenting the report in the particular form adopted.

Part I consists of a non-technical description and explanation of the life table functions, graphs, and other parts of the book. The plan adopted is to give a fairly complete explanation of the various columns contained in the life tables and of the purposes to which they may be put. This is followed by a very extensive set of questions and answers, covering the same ground much more completely and in a way that will be of assistance to many persons because it shows the exact process by which the answers to certain types of questions may be worked out from the mortality tables.

Part II contains the seventy-four life tables which have been constructed in connection with this census investigation. Sixty-four of these tables may be divided into sixteen sets of four each, each set consisting of a table for male lives in 1901, one for males in 1910, one for females in 1901, and one for females in 1910. Seven of the sixteen sets refer to various sections of the population of the original registration states, one set being for the entire population and the others

for whites, negroes, native whites, foreign-born whites, whites in the cities, and whites in rural districts respectively. Five of the sets refer to the population of the individual states of Indiana, Massachusetts, Michigan, New Jersey, and New York, and the remaining four sets to the individual cities of Boston, Chicago, New York, and Philadelphia. Of the remaining ten tables, two are for both sexes in the original registration states in 1901 and 1910 respectively; six, based on the experience of the ten years from 1901 to 1910 inclusive, refer respectively to white males, white females, negro males, and negro females, all in the original registration states, and to negro males and negro females in the District of Columbia; and two represent the experience in 1890 on males and females respectively in the state of Massachusetts. Twenty-five of these tables, constituting all of those based on the mortality rates in 1910 except those for the individual cities just mentioned, were shown in the preliminary report on the United States Life Tables which appeared in 1916.

These very extensive sets of tables render possible many comparisons of mortality in different classes of the community, in different localities and at different times, the comparison in the latter respect indicating the steady improvement in mortality at the younger ages, with indications of a slight change to increased mortality at the older ages.

The life tables dated 1901 are based on the estimated population July 1, 1901, and on the reported deaths in 1900, 1901, and 1902. Those dated 1910 are based on the estimated population July 1, 1910, and on the reported deaths in 1909, 1910, and 1911. Those for the ten years 1901 to 1910 inclusive are based on the estimated mean population for the period and on the aggregate reported deaths. Those for the state of Massachusetts in 1890 are based on the estimated population December 1, 1889, and on the reported deaths in the census year ending May 31, 1890.

In Part III life tables are shown for twelve foreign countries in comparison with those for whites in the original registration states for the years 1901 to 1910 inclusive. These comparisons are shown for male and female lives respectively and for five different columns of the life table, the important ones of course being the annual rate of mortality per thousand and the complete expectation of life in years. There are also shown ten mortality tables based on life insurance experience, five of these tables being based on experience in the United States and five on experience in foreign countries. These are compared with the life table for white males in the original registration states in 1910.

Part IV consists of graphs of various functions of the more important of the life tables now constructed, and of the various population tables and of the life insurance tables shown in comparison therewith.

Part V consists of tables of annuity values, single net premiums, annual net premiums, and commutation columns based on the life tables for white males and white females respectively in 1910. The life annuity values and commutation columns are given for interest at 3, 3½, 4, 5 and 6 per cent, and the net premiums for 3, 3½ and 4 per cent.

Parts VI and VII are devoted to an explanation of the processes adopted in the construction of the life tables, part VI covering the mathematical theory and

Part VII giving an actual illustration of the processes followed in the construction of one of the tables. The table selected as a sample was one based on male lives in the state of New York in 1910.

Part VIII contains the original data upon which all of the tables were based.

The selection of this volume most open to critical analysis is obviously Part VI, which deals with the principles adopted. In the main section of the table running for the principal tables from age 5 to the neighborhood of age 80, the rates of mortality were derived from redistributed figures for population and for deaths derived according to Mr. King's method based on osculatory interpolation. This does not call for any further comment here except on the point that the particular quinquennial age periods selected were those commencing with ages ending in 4 and 9. It is to be noted that Mr. Papps in his recent paper on this subject expresses a preference for periods beginning with ages ending in 3 and 8, whereas the method adopted by Mr. King for the English Life Table No. 8 agrees with that adopted for the United States Life Tables.

The rates of mortality at the extreme old ages were derived from the application of Wittstein's formula, the constants being determined from rates of mortality at ten consecutive ages derived according to the method proposed by Mr. King to obtain graduated pivotal values for quinquennial interpolations. These rates of mortality are also used to supply preliminary values for a group of ages just below those to which Wittstein's formula is applied. This whole section of the table is finally graduated by Spencer's 21 term formula, and the rates thus arrived at are used to give a gradual transition from those in the middle section of the table to those derived by Wittstein's formula.

The only criticism which I would make of this part of the processes is that the means adopted to weld the two parts of the table together seem to be rather more elaborate than should be necessary. Perhaps, however, they are more complicated in explanation than in actual application.

Possibly the most interesting part of the explanation is that relating to the adjustment of the population at infantile ages. An examination of the population returns by individual ages indicates very clearly a deficiency of population at the two youngest ages, and a comparison with the birth returns indicates that this is not a transfer to other ages but an absolute failure to report. Some method was therefore necessary to estimate the true population at these ages. The method described in this volume may be summarized as an estimate of the population at ages 2 to 5 two years after the close of the period of observation, or three and one-half years after the middle of that period, combined with an analysis of the death returns of the period of observation and the succeeding two calendar years so as to determine, approximately at least, how many of these deaths should be assigned to those born during the three calendar years of observation. It was then assumed that immigration could be ignored at these ages and that the number of births during the period of observation would be obtained by adding these deaths to the estimated survivors two years after the end of the period, their ages being then between 2 and 5 years.

An estimate was then made of the increase in the population at each individual age during the three years, and by successively subtracting from the number of

births these increases and the deaths and making the proper adjustment, the number exposed to risk was obtained for the individual ages. The exposed to risk thus obtained was used for all ages up to four, last birthday inclusive.

It will be noted that these recalculated exposures were used for ages 2, 3, and 4 last birthday, although statistics were available for the population at those ages which were considered as correct, at least in the aggregate, and were used as the basis of the estimate. This suggests a comparison between the average population at these ages as derived from the recalculated exposed risk and the original statistics. This comparison for male lives in the state of New York is shown below, together with other redistributions of the population which will be referred to later.

It will be noted that the total population at these three ages according to the official adjustment is 2,601 less than the total population at the same ages according to the original statistics, which fact indicates an average increase of nearly 1 per cent in the rate of mortality at these ages and suggests the idea that possibly more satisfactory results would have been obtained by sub-

AVERAGE POPULATION

Ages	Original statistics	Official adjustment	First correction	Second correction
2-3	94,229	92,372	93,239	93,186
3-4	90,718	89,476	90,343	90,454
4-5	86,876	87,374	88,241	88,183
	271,823	269,222	271,823	271,823

stituting for the method actually used in estimating the births during the period one based on the idea of so determining these births that after the rest of the process had been followed through, the total of the population at ages 2 to 5 would agree with the original statistics. It is easy to see that this correction would amount to adding 2,601 to the estimated births during the period and consequently to the exposed to risk at each subsequent age. This would amount to adding 867 to the average population at each age and would give the figures contained in the foregoing table under the heading "First correction." This adjustment although probably sufficiently satisfactory for all practical purposes is still liable to the criticism that although the population according to the original statistics for ages 0 to 1 last birthday is deficient, it is used as the basis for the estimate of the increase in population at those ages during the period of observation, and that while we have considered the population of the ages between 2 and 5 as subject to redistribution, varying percentages of increase are used for these ages determined from the original statistics. This objection could be obviated by using a uniform percentage under 5, based on the combined statistics for ages between 2 and 5 and deriving a recurrence formula based on the principle that the increase in population at each age bears a fixed ratio to the finally adjusted population rather than to the original statistics. The figures for the population given in the above table under the heading "Second correction" are derived according

to this method, and the following table gives the rates of mortality at ages 0 to 4 inclusive according to the original statistics unadjusted and according to the various adjustments above referred to.

MALE LIVES IN THE STATE OF NEW YORK

Rates of Mortality per 1000

Age	Unadjusted	Official	First correction	Second correction
0	135.31	127.13	126.18	126.38
1	37.32	33.64	33.35	33.37
2	15.25	15.56	15.42	15.42
3	9.18	9.31	9.22	9.20
4	6.59	6.55	6.49	6.49

An investigation of the reason for the deficiency in population, disclosed by the official readjustment, indicates that this is principally due to the use of actual death returns for the years 1912 and 1913, which not only did not increase in proportion to the estimated increased population but actually decreased. The actual deaths added in arriving at the assumed births according to the official adjustment were 57,529, and the corresponding deaths during the three years 1909 to 1911 inclusive were 58,926, showing a decrease of 1,397 instead of the expected increase.

The whole volume forms a very valuable addition to our records regarding mortality experience, and it is well worth the study of those interested in the subjects covered by it.

ROBERT HENDERSON

Introduction to Economic Statistics, by George R. Davies. New York; The Century Company. 1922. vi, 163 pp.

The economist who would employ quantitative methods in attacking his problems is enabled to do so largely because of contributions made by workers in other fields. Methods of dealing with assembled measurements and with mass phenomena which have been developed by the biometrician, the demographer, the astronomer, the mathematician, the logician, have been taken over by the economist. Though many economists—Jevons, Edgeworth, Moore—have contributed to the development of the science, statistics is to economics an adopted child or, to change the figure, an alien who must be naturalized. The methods of the biometrician cannot be employed unchanged by the economist. Adaptations must be made, new devices employed, and tests made of the validity of old methods when applied to the raw materials of economics. All this is being done. Methods borrowed from Quetelet, Galton, Pearson, and the other fathers of modern statistics are being welded into a body of method designed for handling the data of economics. This by no means involves the building up of a new science of statistics, but merely modification and adaptation, with the central core of the science unchanged.

The fact that the science of statistics is largely of alien birth has been something of a barrier to the adoption of quantitative methods in economics. The subject has been expounded, to a great extent, in terms of anthropometry, demography, education. Leaves, nuts, clam-shells, and cricket batting averages have bulked large in textbook and classroom. The terminology—lines of regression, probable error—bears the imprint of an alien origin. The economist is deeply indebted to those from whom quantitative methods have been borrowed, but the need has been pressing for the presentation of statistical methods in a form better adapted to his immediate needs, with dress and mannerisms which will not evoke the suspicion which adheres to strange ways. Within the last few years several books have appeared in this country which help to meet this need. The book at present under review is the most recent of these publications.

Within the 132 pages of text there is treatment of the subjects of tabulation, types and measures of dispersion, price, wage and quantity index numbers, trends and cycles, and correlation, with an appendix dealing with the laboratory material required for ordinary statistical and graphic work. A set of very useful laboratory exercises follows each of the six chapters. In the sections on index numbers and on trends and cycles the writer is dealing with subjects in which the economist and business statistician are vitally interested, and which have not been covered adequately in most of the standard works on statistics. The methods of analyzing and comparing time series discussed in these chapters rest in part upon recent work by the author and other statisticians, and this explanation constitutes a welcome addition to existing material. The exposition is concise and lucid throughout. Illustrative examples are numerous and well-chosen.

Condensed and clear as the treatment is, the author fails to cover completely the field marked out by his six chapter headings. The treatment of the fundamentals of statistical method is inadequate, if the book is to stand by itself as an exposition of the subject. One whose knowledge of statistics was based upon this book alone would have an acquaintance with certain methods, but an incomplete understanding of the basic reasoning and the assumptions lying back of these methods. It is not only that underlying hypotheses are undeveloped: the explanation of the statistical measures and methods dealt with in the book is in some cases too condensed for the beginner, for whom the book is avowedly written. It is inevitable that this should be so in such a brief treatment, and the author undoubtedly recognizes this limitation. When supported by classroom explanation and laboratory work these difficulties would not be experienced, and it is from such a situation that the book developed. (The book is an outgrowth of an undergraduate course in statistics given at Princeton University during the academic year 1920–21.) But when it stands alone these inadequacies would be felt, and consultation of other books, to which the reader is introduced by numerous references, would be necessary.

But these criticisms are perhaps ungracious. Here is a useful contribution, a well-written presentation of the statistical methods which the economist must employ, prepared with the needs of the economist in mind. The alien atmosphere of mortality statistics, cephalic indexes, and lobster lengths has been dis-

sipated. All who are interested in economic study must be grateful for the impetus which is given by such works as this to the development of quantitative methods of research in the field of economics, to the collection and analysis of hard facts in studying the working of the business system.

FREDERICK C. MILLS

Columbia University

The Clothing Workers of Chicago, 1910–1922, by Leo Wolman, Eleanor Mack, H. K. Herwitz, and Paul Wander, Chicago: The Chicago Joint Board, Amalgamated Clothing Workers of America, 1922. 424 pp.

There is an ideal economic world, now forsaken by most economists but still assiduously preached by newspaper editorial writers, in which immutable laws operate with automatic precision to defeat the human will, especially when the will is expressed in group action. In that world wages and prices are fixed solely by competitive forces under the conditions of demand and supply. Labor, capital, land, and the entrepreneur each receive what they respectively contribute to production—no less and no more. There are unalterable relations between prices and wages: when wages go up, the wage-earners are inevitably defeated by corresponding rises in prices, and when prices go down, wages must be equally "deflated." In that world, labor organization, in so far as it attempts to wield economic power, is both useless and pernicious. The necessary exponent of economic law is the individual employer; through him, as through a king enthroned in divine right, Destiny is made manifest. The Open Shop, regulated by the Efficiency Engineer, is the loyal and happy kingdom.

Most exponents of this doctrine hold the delusion that the ideal is real. They exhort people to be resigned and not to kick against the pricks. But people insist on kicking. The reality is full of unions and strikes. Industrial potentates are defied; their powers are limited by forces other than economic law. All this goes on in prosperity and in depression, despite exhortation.

The Clothing Workers of Chicago, 1910–1922, is significant because it is a careful picture, not of the ideal world of automatic economics, but of a part of the real world in which the organized wage-earners have fought their way to recognition and have succeeded in expressing a collective will in an industry. Here the dogma of divine right has been frankly abandoned, and the democratic experiment of collective regulation has won its preliminary fight for life and been given a chance to try what it can do through constitutional government.

In the beginning the clothing industry was a paradise according to the theory of free competition. No industry was more competitive, none was less subject to any regulation outside of "economic law." The result was not a paradise for the workers. It was embodied in the sweatshop, in starvation wages, in unlimited hours of work at some periods of the year and unemployment at others. The workers rebelled. The early history of that rebellion shows all the marks of the type of industrial conflict now going on in many other industries, which is ordinarily assumed to be futile. It includes bitter and lost strikes, short-sighted

union policy, even grafting union officials. It includes injunctions, defiance from self-righteous employers, "radicalism," and all the rest. But eventually the workers won a foothold in the shop of Hart, Schaffner and Marx, and the machinery of government began to develop. It now embraces the whole of the Chicago market, with its 40,000 to 50,000 employees, and has brought comparative industrial peace—as well as higher wages, shorter hours, and better conditions throughout.

This history of the union's struggle for recognition, and of the subsequent development of constitutional government in industry, prepared under the direction of Mr. Leo Wolman with the coöperation of Miss Eleanor Mack, Mr. H. K. Herwitz and Mr. Paul Wander, gives an impressive picture of collective action which cannot be ignored by students of economics. It explains, in the first place, how the collective economic power of the workers brought about a system of adjustment having two main features: First, a machinery of general arbitration, and second, a machinery for the application of general principles to specific cases through a Trade Board and an Impartial Chairman. It explains the structure of the union, showing how it furnishes the cohesive force in the industry through the Joint Board and the subordinate shop chairmen. Perhaps the most important of all, however, is the section devoted to an analysis of the decisions and precedents arrived at in the operation of this industrial government in important matters such as the powers of management, discipline and discharge, the introduction of new processes and machinery, the adjustment of wages, and the principle of union preference (an issue usually inaccurately stated in terms of the "open" *vs.* the "closed" shop).

The book should be sharply differentiated from the many current treatises which approach the "labor problem" from the point of view of a type of management which may allay or prevent open discontent, or from the point of view of arbitration or adjustment imposed from above chiefly with a view to protecting "the public" from inconvenience or loss. It is an invaluable study, not of a patent scheme, but of a genuine social and economic development in industry arising chiefly from the workers themselves.

GEORGE SOULE

Annual Report on the Demography and Public Health of Brussels for the Year 1920.

This admirable publication reflects throughout with characteristic Belgian thoroughness the gratifying tendency of that nation toward the road of social and economic recovery. The enumerated population of Brussels as of January 1, 1920, is given as 156,924, of which 71,134 were males and 85,790 were females. Of the total population only 62 were 90 years of age or over, while the proportion at ages under 1 was 1.32 per cent. The density of population for the city of Brussels is 146 per hectare, but it runs as high as 432 for the second territorial division of the city. The birth rate for the year 1920 was 17.9 per 1,000, which is a gratifying increase over the average for 1911–19, given as 13.4, but it is far below the average for the period 1867–70, when the rate was as high as 38.8. The statistics of illegitimacy indicate a lamentable moral condition, the percentage

of illegitimate births for 1920 having been 24.0, which, however, is somewhat less than the average for 1911–19, given as 28.0. Equally disturbing is the high number of divorces, which is given as 171 for the year 1920, which compares with an average of 114 for the period 1911–19.

The death rate during 1920 was 13.0 per 1,000, which contrasts with a death rate of 29.8 during 1867–70 and 16.7 during 1911–15. Contagious diseases have diminished very materially during recent years, including the war period, the rate having declined from 4.4 per 1,000 in 1867–70 to 0.7 in 1901–10, 0.4 during 1911–15, 0.3 during 1916–19, and only 0.2 during 1920.

There was an astonishing decline in the mortality from pulmonary tuberculosis, the lowest figure on record having been reached during 1920, when the rate was only 9.7 per 10,000, in contrast to a rate of 49.4 during 1867–70. During the early war period the rate was 16.2, and during the later period, 23.8 per 10,000 of population.

In contrast, there was a very slight increase in the mortality from cancer in 1920 as compared with 1919, the rate for the later year being lower than during any previous year except 1914 and 1916.

The actual deaths from pulmonary tuberculosis decreased from 267 during 1919 to 152 during 1920. Among other interesting facts disclosed by the table of causes in detail reference may be made to the mortality from suicide, which during 1920 decreased to 39 as against 45 in 1919; and there were only two deaths from homicide during 1920 in a population, as previously stated, of 156,924. During the previous year the number of deaths from homicide was 4. There was only one death from typhoid fever, five from smallpox, two from scarlet fever, and ten from whooping cough.

These observations apply exclusively to the city of Brussels and have no reference to the suburbs, the population of which increases the number of inhabitants of Greater Brussels to 811,690. For this area the death rate during 1920 was 13.3, while the birth rate was 17.4, or about the same as for the city of Brussels in the more restricted sense of the term. The report is illustrated by a number of admirable diagrams and maps.

F. L. HOFFMAN

The Mathematical Theory of Probabilities and Its Application to Frequency Curves and Statistical Methods, by Arne Fisher. *Volume I. Mathematical Probabilities, Frequency Curves, Homograde and Heterograde Statistics.* 2d. ed., greatly enlarged. New York: The Macmillan Company. 1922. xxix, 289 pp.

The enlargement of the second edition of this volume consists in adding to the 171 pages of the old edition, which now appears as Part I, two new sections: Part II, pp. 169–214, on Frequency Curves and Heterograde Statistics, and Part III, pp. 215–89, on Practical Applications of the Theory. It is therefore practicable to let the earlier review[1] stand as written and turn attention to the additions.

[1] This JOURNAL, Vol. VI, Dec. 1916, pp. 468–9.

In Part II there are two chapters: XIII, The Theory of Errors and Frequency Curves and Its Application to Statistical Series. General Remarks, p. 169; XIV, The Mathematical Theory of Frequency Curves, p. 188. In Part III are four chapters: XV, The Numerical Determination of the Parameters, p. 215; XVI, Logarithmically Transformed Frequency Functions, p. 235; XVII, Frequency Curves and their Relation to the Bernoullian Series, p. 261; XVIII, Poisson-Charlier Frequency Curves for Integral Variates, p. 271. These titles indicate in a fairly clear manner the contents of those new sections of the volume which follow and depend upon the general theory of probabilities. Parts II and III are not only more advanced, they are decidedly harder reading than Part I. The reader who is unfamiliar with mathematics will find little cheer. A well-known quotation comes to mind: Je me rends parfaitement compte du désagréable effet que produit sur la majorité de l'humanité, tout ce qui se rapporte, même au plus faible dégré, à des calculs on raisonnements mathématiques. Still, there is a point beyond which one cannot go in statistics unless he is able to overcome his dislike for mathematical analysis or at least for mathematical reasoning.

In this book Mr. Fisher shows that he belongs to the school of Laplace, Poisson, Gram, Charlier—the French-Scandinavian school rather than the German or the English. This is fortunate, for it brings into our scientific literature a presentation of ideas hitherto almost if not quite lacking in it. We are promised by Mr. Fisher a continuing volume which will be eagerly awaited, particularly as it will contain not only additional statistical theory and application but mathematical tables advantageous to those who would apply the theory already incorporated in the new Volume I.

EDWIN BIDWELL WILSON

Harvard School of Public Health

RECENT LITERATURE

STATISTICAL BOOKS AND PUBLICATIONS

Annual report of the Commissioner of Mental Diseases, Massachusetts, 1920. Public document, No. 117. 264 pp.

Austrian, J. E. Digest of business statistics. New York: J. E. Austrian. 1922.

Bartley, Joseph C. A study of price control by the United States Food Administration. A dissertation submitted to the Faculty of Philosophy of the Catholic University of America. Villanova, Pa. 1922.

Bureau of Business Research, Northwestern University School of Commerce. Horace Secrist: The control of expenses and profits in the retail merchandising of clothing. Chicago. 1922. 12 pp. (Series II, No. 6.)

——Horace Secrist and Blanch L. Altman: Amounts sold per full-time salesmen in retail clothing stores. Chicago. 1922. 14 pp. (Series II, No. 5.)

——James Grantham: Wholesale price movements of paper in Chicago, January 1, 1913, to June 30, 1922. Chicago. 1922. 13 pp. (Series I, No. 2.)

Cassel, Gustav. Money and foreign exchange after 1914. New York: Macmillan Company. 1922.

Hirst, Francis W. The paper moneys of Europe. New York: Houghton Mifflin Co. 1922.

International Labour Office. Studies and Reports, 1922.

No. 1. Statistics of unemployment in various countries, 1910 to 1922. Geneva. 1922. 28 pp.

——No. 2. Wage changes in various countries, 1914 to 1921. Geneva. 1922. 76 pp.

League of Nations, Health Section. No. 3. Epidemiological intelligence. Geneva. 1922. 43 pp.

March, Lucien. Report of the International Statistical Commission of the League of Nations. 1920. 35 pp.

National Bureau of Economic Research. The Income in the United States. Vol. II. New York: Harcourt, Brace and Co. 1922.

Pennsylvania Department of Internal Affairs. Report on productive industries, railways, taxes and assessments, waterways, and miscellaneous statistics for 1920. Harrisburg, Pa. J. L. Kuhn. 1921. 1,040 pp.

Reed, H. L. The development of federal reserve policy. New York: Houghton Mifflin Co. 1922.

Unemployment in East London. The report of a survey made from Toynbee Hall. London: P. S. King and Son. 1922.

GOVERNMENT PUBLICATIONS

United States. Bureau of Education. Bulletin No. 37. Malnutrition and school feeding, by J. C. Gebhart. 39 pp.

——Bureau of Labor Statistics. Bulletins.

No. 296. Wholesale prices, 1890–1920. 247 pp.

No. 302. Union scale of wages and hours of labor, May 15, 1921. 219 pp.

No. 304. Proceedings of the Eighth Annual Meeting of the International Association of Industrial Accident Boards and Commissions, 1921. 254 pp.

No. 311. Proceedings of the Ninth Annual Meeting of the International Association of Public Employment Services, 1921. 78 pp.

No. 316. Hours and earnings in anthracite and bituminous coal mining. Anthracite, January, 1922. Bituminous, winter of 1921–22. 63 pp.

——Children's Bureau. Bulletin No. 87. Statures and weights of children under six years of age, by R. M. Woodbury. 117 pp.

——Decisions of the United States Labor Board, Vol. II. 1921. 1,007 pp.

——Public Health Bulletins.

No. 116. Lead poisoning in the pottery trades, by B. J. Newman, W. J. McConnell, O. M. Spencer, and F. M. Phillips. 223 pp.

No. 124. The future of public health in the United States and the education of sanitarians. 42 pp.

——United States Tariff Commission. Colonial tariff policies. 869 pp.

Belgium. Annuaire statistique de la Belgique et du Congo Belge, 1915–1919, Brussells. 1922. 240 pp.

Czecho-Slovakia. Bulletin statistique de la République Tchécoslovaque. 1922.

France. Annuaire statistique. Vol. 37. Paris. 1922. 449 pp.

——Bulletin de la Statistique Générale de la France, July, 1922.

Germany. Statistiches Jahrbuch der Stadt Berlin, 1915–1919. Berlin. 1922. 935 pp.

ARTICLES IN PERIODICALS

Casualty Actuarial Society. Proceedings. May, 1922. A. H. Mowbray: The value of a social point of view in the conduct of the casualty insurance business. V. N. Valgren: Agricultural insurance. E. H. Downey: Remarriage experience of Pennsylvania compensation insurance carriers policy, years 1916–19. L. I. Dublin and E. W. Kopf: Mortality from external causes among industrial policyholders of the Metropolitan Life Insurance Company, 1911–20. A. H. Mowbray: Observations on pension funds for employees rendered permanently disabled by reason of a second injury. J. E. Gregory: Credit insurance.

International Labour Review. Vol. VI.

September, 1922. Leymann: The eight-hour day and the problem of overtime in Germany. Albert Mansbridge: The workers' educational association of Great Britain. New agrarian legislation in Central Europe. Unemployment insurance: an international survey.

October, 1922. Fritz Sitzler: The law of collective bargaining in Germany. Henri de Man: Workers' education in Belgium. The three-shift system in the iron and steel industry.

Journal of Political Economy.

August, 1922. J. A. Hobson: Britain's economic outlook on Europe. H. H. Preston: Branch banking in California. A. H. Hansen: The economics of unionism. W. H. George: Proudhon and economic federalism. J. W. Thompson: Early German-Slav trade. A. C. Hodge: Bases of control for retail inventory. H. G. Moulton: The Danner-Kraft Dry Goods Company.

October, 1922. R. Lennard: English agriculture since 1914. I. A. Bergland: Iron and shell and Japanese policies. Jacob Viner: Dumping in international trade. I. C. Southworth: Newsprint paper industry and the tariff. E. S. Gregg: Shipping fallacies. J. F. Pyle: The taxation of incomes in Oklahoma.

Journal of the Indian Economic Society.

June, 1922. C. S. Deole: Indian military expenditure. S. N. Pherwani: Economics of education. C. N. Vakil: Famine insurance grant.

March, 1922. Indian budget for 1922–23. C. S. Deole: Economics of Indian railways. R. M. Joshi: Welfare work, trade unionism, and cöoperation.

Journal of the Institute of Actuaries. No. 278. July, 1922.

G. King: On a short method of constructing select mortality tables. C. H. Maltby: Results of an investigation into the effect of different valuation bases upon surplus. I. F. Steffensen: On certain formulas of approximate summation and integration.

Monthly Labor Review. June–October, 1922.

June. Margaret Gadsby: Development of collective bargaining in the men's clothing industry in the United States.

July. Ethelbert Stewart: Efficiency of American labor. J. P. Meade: Problems and importance of factory inspection.

August. W. W. Husband: Immigration under the per centum limit act. Ethelbert Stewart: "1913=100."

September. M. G. Lloyd: Preparation of safety codes under the auspices of the American Engineering Standards Committee. John Roach: Factory chiefs in New Jersey.

October. M. G. Dill and E. E. Witte: State legislation upon private employment agencies.

OTHER BOOKS AND PUBLICATIONS

Devine, Edward T. Social work. New York: Macmillan Company. 1922. vii, 352 pp.

Lauck, W. Jett, and Watts, Claude S. The industrial code. New York: Funk and Wagnalls Company. 1922. v, 571 pp.

Memorias y Revista de la Sociedad Cientifica "Antonio Alzate." Mexico. June, 1922.

Ogburn, W. F. Social Change. New York: B. W. Huebsch, Inc. 1922.

Proceedings of the Royal Society of Edinburgh, Session 1921–22.